

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**Regulatory DNA sequences of the gene for the human catalytic telomerase subunit, and their diagnostic and therapeutic use**

**Structure and function of the chromosome ends**

5

The genetic material of eukaryotic cells is distributed on linear chromosomes. The ends of hereditary units are termed telomeres, derived from the Greek words *telos* (end) and *meros* (part, segment). Most telomeres consist of repeats of short sequences which are mainly composed of thymine and guanine (Zakian, 1995). In all the vertebrates which have so far been investigated, the telomeres consist of the sequence TTAGGG (Meyne *et al.*, 1989).

10

15

The telomeres have a variety of important functions. They prevent the fusion of chromosomes (McClintock, 1941) and thus the formation of dicentric hereditary units. Such chromosomes having two centromeres can lead to the development of cancer due to loss of heterozygosis or duplication, or loss of genes.

20

In addition, telomeres serve the purpose of distinguishing intact hereditary units from damaged hereditary units. Thus, yeast cells ceased their cell division when they contained a chromosome without a telomere (Sandell and Zakian, 1993).

25

30

Telomeres fulfil another important task in association with the replication of eukaryotic cell DNA. In contrast to the circular genomes of prokaryotes, the linear chromosomes of eukaryotes cannot be completely replicated by the DNA polymerase complex. RNA primers are required to initiate DNA replication. After elimination of the RNA primers, extension of the Okazaki fragments and subsequent ligation, the newly synthesized DNA strand lacks the 5' end since the RNA primer cannot be replaced by DNA at that point. Without special protective mechanisms, the chromosomes would therefore shrink with each cell division ("end-replication problem"; Harley *et al.*, 1990). The non-coding telomere sequences presumably constitute a buffer zone for preventing the loss of genes (Sandell and Zakian, 1993).

001260 "Sheets"

In addition to this, telomeres also play an import role in regulating cell ageing (Olovnikov, 1973). Human somatic cells exhibit a limited capacity for replication in culture; after a certain period of time, they become senescent. In this state, the cells no longer divide even after having been stimulated with growth factors; however, they do not die and remain metabolically active (Goldstein, 1990). Various observations support the hypothesis that a cell determines how many more times it can divide on the basis of the length of its telomeres (Allsopp *et al.*, 1992).

In summary, the telomeres consequently possess key functions in the ageing of cells, and in stabilizing the genetic material and preventing cancer.

#### The enzyme telomerase synthesizes the telomeres

As described above, organisms which possess linear chromosomes can only replicate their genome incompletely in the absence of a special protective mechanism. Most eukaryotes use a special enzyme, i.e. telomerase, for regenerating the telomere sequences. Telomerase is expressed constitutively in the single-cell organisms which have so far been investigated. On the other hand, telomerase activity has only been measured in humans in germ cells and tumour cells, whereas neighbouring somatic tissue did not contain any telomerase (Kim *et al.*, 1994).

Telomerase can also be designated functionally as terminal telomere transferase, which is located in the cell nucleus as a multiprotein complex. While the RNA moiety of human telomerase has been known for a relatively long period of time (Feng *et al.*, 1995), the catalytic subunit of this enzyme group was recently identified in a variety of organisms (Lingner *et al.*, 1997; cf. our application PCT EP/98/03468 which is likewise pending). These catalytic subunits of telomerase are strikingly homologous both among themselves and in relation to all previously known reverse transcriptases.

WO 98/14592 also describes nucleic acid and amino acid sequences of the catalytic telomerase subunit.

Activation of telomerase in human tumours

It was originally only possible to demonstrate telomerase activity in humans in germ  
line cells and not in normal somatic cells (Hastie *et al.*, 1990; Kim *et al.*, 1994).  
Following the development of a more sensitive detection method (Kim *et al.*, 1994),  
a low telomerase activity was also detected in hematopoietic cells (Broccoli *et al.*,  
1995; Counter *et al.*, 1995; Hiyama *et al.*, 1995). It is true, however, that these cells  
nevertheless exhibited a reduction in the telomeres (Vaziri *et al.*, 1994; Counter *et al.*,  
1995). It has still not been resolved whether the quantity of enzyme in these cells  
is not sufficient for compensating the telomere loss or whether the telomerase activity  
which is measured stems from a subpopulation, e.g. incompletely differentiated  
CD34<sup>+</sup>38<sup>+</sup> precursor cells (Hiyama *et al.*, 1995). In order to resolve this, it would be  
necessary to detect telomerase activity in a single cell.

Interestingly, however, significant telomerase activity was detected in a large number  
of the tumour tissues which had thus far been tested (1734/2031, 85%; Shay, 1997),  
whereas no activity was found in normal somatic tissue (1/196, <1%, Shay, 1997). In  
addition various investigations have shown that the telomeres still shrank in  
senescent cells which were transformed with viral oncoproteins and it was only  
possible to detect telomerase in the subpopulation which survived the growth crisis  
(Counter *et al.*, 1992). The telomeres were also stable in these immortalized cells.  
(Counter *et al.*, 1992). Similar findings from investigations in mice (Blasco *et al.*,  
1996) support the assumption that reactivation of the telomerase is a late event in  
tumorigenesis.

Based on these results, a "telomerase hypothesis" was developed which links the loss  
of telomere sequences and cell ageing with telomerase activity and the development  
of cancer. In long-lived species such as humans, the shrinking of the telomeres can be  
regarded as being a mechanism for suppressing tumours. Differentiated cells which  
do not contain any telomerase cease their cell division at a particular telomere length.  
If such a cell mutates, it can only form a tumour if the cell can extend its telomeres.

Otherwise, the cell would continue to lose telomere sequences until its chromosomes became unstable and it was finally destroyed. Telomerase reactivation is presumably the main mechanism used by tumour cells to stabilize their telomeres.

5 It follows from these observations and considerations that it should be possible to treat tumours by inhibiting the telomerase. Conventional cancer therapies using cytostatic agents or short-wave radiation damage all the dividing cells in the body in addition to the tumour cells. However, since only germ line cells, apart from tumour cells, contain significant telomerase activity, telomerase inhibitors would attack the  
10 tumour cells more specifically and consequently elicit fewer undesirable side effects. Telomerase activity has been detected in all the tumour tissues which have so far been tested, which means that these therapeutic agents could be employed against all types of cancer. The effect of telomerase inhibitors would then set in when the telomeres of the cells had shortened to such an extent that the genome became  
15 unstable. Since tumour cells usually possess telomeres which are shorter than those of normal somatic cells, cancer cells would be the first to be eliminated by the telomerase inhibitors. By contrast, cells possessing long telomeres, such as the germ cells, would only be damaged at a much later date. Telomerase inhibitors consequently represent a potential way forward in the treatment of cancer.

20

It becomes possible to obtain unambiguous answers to the question of the nature and points of attack of physiological telomerase inhibitors once the manner in which expression of the telomerase gene is regulated has also been identified.

25 Regulation of gene expression in eukaryotes

There are a large number of points in eukaryotic gene expression, i.e. the cellular flow of information from the DNA to the protein by way of the RNA, at which regulatory mechanisms can exert an effect. Examples of individual control steps are  
30 gene amplification, the recombination of gene loci, chromatin structure, DNA methylation, transcription, post-transcriptional modifications of mRNA, mRNA transport, translation and post-translational modifications of proteins. Studies which

have been carried out to date indicate that control at the level of transcription initiation is of the greatest importance (Latchman, 1991).

5 A region which is responsible for regulating transcription, and which is designated the promoter region, is located directly upstream of the transcription start of a gene which is transcribed by RNA polymerase II. Comparison of the nucleotide sequences of promoter regions from a large number of known genes shows that particular sequence motifs occur regularly in this region. These elements include, inter alia, the TATA box, the CCAAT box and the GC box, which elements are recognized by  
10 specific proteins. The TATA box, which is located about 30 nucleotides upstream of the transcription start, is, for example, recognized by the TFIID subunit TBP ("TATA box-binding protein"), whereas particular GC-rich sequence segments are specifically bound by the transcription factor Sp1 ("specificity protein 1").

15 The promoter can be functionally subdivided into a regulatory segment and a constitutive segment (Latchman, 1991). The constitutive control region comprises the so-called core promoter which enables transcription to be initiated correctly. This promoter contains the sequence elements which are described as UPE's (upstream promoter elements) which are necessary for efficient transcription. The regulatory  
20 control segments, which can be interlaced with the UPE's, possess sequence elements which can be involved in the signal-dependent regulation of transcription by hormones, growth factors, etc. They impart tissue-specific or cell-specific promoter properties.

25 DNA segments which are able to exert an influence on gene expression over relatively large distances are a characteristic feature of eukaryotic genes. These elements can be located upstream or downstream of a transcription unit, or within the unit, and can perform their function independently of their orientation. These sequence segments may reinforce (enhancers) or attenuate (silencers) promoter  
30 activity. In a similar way to the promoter regions, enhancers and silencers also accommodate several binding sites for transcription factors.

5 The invention particularly relates to the 5'-flanking regulatory DNA sequence which contains the promoter DNA sequence for the gene for the human catalytic telomerase subunit, as depicted in Fig. 10 (SEQ ID NO 3).

10

15

20

25

30

Particular preference is given to those antineoplastic proteins which inhibit angiogenesis directly or indirectly. Examples of these proteins are:

5 Plasminogen activator inhibitor (PAI-1), PAI-2, PAI-3, angiostatin, endostatin, platelet factor 4, TIMP-1, TIMP-2, TIMP-3 and leukaemia inhibitory factor (LIF).

Antineoplastic proteins which have a direct or indirect cytostatic effect on tumours are likewise particularly preferred. These proteins include, in particular:

10 perforin, granzyme, IL-2, IL-4, IL-12, interferons, such as IFN- $\alpha$ , IFN- $\beta$  and IFN- $\gamma$ , TNF, TNF- $\alpha$ , TNF- $\beta$ , oncostatin M; tumour suppressor genes, such as p53, retinoblastoma.

15 Particular preference is furthermore given to antineoplastic proteins which, where appropriate in addition to their antineoplastic effect, stimulate inflammations and thereby contribute to the elimination of tumour cells. Examples of these proteins are:

20 RANTES, monocyte chemotactic and activating factor (MCAF), IL-8, macrophage inflammatory protein (MIP-1 $\alpha$ , - $\beta$ ), neutrophil activating protein-2 (NAP-2), IL-3, IL-5, human leukaemia inhibitory factor (LIF), IL-7, IL-11, IL-13, GM-CSF, G-CSF and M-CSF.

25 Particular preference is furthermore given to antineoplastic proteins which, due to their action as enzymes, are able to convert precursors of an antineoplastic active compound into an antineoplastic active compound. Examples of these enzymes are:

30 herpes simplex virus thymidine kinase, varicella zoster virus thymidine kinase, bacterial nitroreductase, bacterial  $\beta$ -glucuronidase, plant  $\beta$ -glucuronidase from *Secale cereale*, human glucuronidase, human carboxypeptidase, bacterial carboxypeptidase, bacterial  $\beta$ -lactamase, bacterial cytosine deaminidase, human catalase and/or phosphatase, human alkaline phosphatase, type 5 acid phosphatase, human



lysooxidase, human acid D-aminooxidase, human glutathione peroxidase, human eosinophil peroxidase and human thyroid peroxidase.

5 The abovementioned recombinant constructs can also contain DNA sequences which encode factor VIII or factor IX, or part fragments thereof. These DNA sequences also include other blood clotting factors.

The abovementioned recombinant constructs can also contain DNA sequences which encode a reporter protein. Examples of these reporter proteins are:

10 Chloramphenicol acetyl transferase (CAT), glow-worm luciferase (LUC),  $\beta$ -galactosidase ( $\beta$ -Gal), secreted alkaline phosphatase (SEAP), human growth hormone (hGH),  $\beta$ -glucuronidase (GUS), green-fluorescing protein (GFP), and all the variants derived therefrom, aquarin and obelin.

15 Recombinant constructs according to the invention can also contain DNA which encodes the human catalytic telomerase subunit and its variants and fragments in the antisense orientation. Where appropriate, these constructs can also contain other protein subunits of the human telomerase and the telomerase RNA component in the  
20 antisense orientation.

The recombinant constructs can, in addition to the DNA which encodes the human catalytic telomerase subunit, and its variants and fragments, also contain other protein subunits of the human telomerase and the telomerase RNA component.

25 The invention furthermore relates to a vector which contains the abovementioned DNA sequences according to the invention, in particular the 5'-flanking DNA sequences and also one or more of the other DNA sequences mentioned above.

30 The preferred vector for these constructs is a virus, for example a retrovirus, an adenovirus, an adeno-associated virus, a herpes simplex virus, a vaccina virus, a lentiviral virus, a Sindbis virus and a Semliki forest virus.

Preference is also given to using plasmids as vectors.

5 The invention furthermore relates to pharmaceutical preparations which comprise recombinant constructs or vectors according to the invention; for example a preparation in a colloidal dispersion system.

Examples of suitable colloidal dispersion systems are liposomes or polylysine ligands.

10

The preparations of the constructs or vectors according to the invention in colloidal dispersion systems can be supplemented with a ligand which binds to the membrane structures of tumour cells. Such a ligand can, for example, be attached to the construct or the vector or else be a component of the liposome structure.

15

Suitable ligands are, in particular, polyclonal or monoclonal antibodies, or antibody fragments thereof, which bind, by their variable domains, to the membrane structures of tumour cells, or substances carrying mannose terminally, cytokines or growth factors, or fragments or part sequences thereof, which bind to receptors on tumour cells.

20

Examples of corresponding membrane structures are receptors for a cytokine or a growth factor, such as IL-1, EGF, PDGF, VEGF, TGF  $\beta$ , insulin or insulin-like growth factor (ILGF), or adhesion molecules, such as SLeX, LFA-1, MAC-1, LECAM-1 or VLA-4, or the mannose-6-phosphate receptor.

25

The present invention includes pharmaceutical preparations which, in addition to the vector constructs according to the invention, can also comprise non-toxic, inert, pharmaceutically suitable excipients. It is possible to conceive of administering (e.g. intravenously, intraarterially, intramuscularly, subcutaneously, intradermally, anally, vaginally, nasally, transdermally, intraperitoneally, as an aerosol or orally) these preparations at the site of a tumour or administering them systemically.

30

The vector constructs according to the invention can be employed in gene therapy.

5 The invention furthermore relates to a recombinant host cell, in particular a recombinant eukaryotic host cell, which harbours the above-described constructs or vectors.

10 The invention furthermore relates to a process for identifying substances which affect the promoter activity, silencer activity or enhancer activity of the catalytic telomerase subunit, with this process comprising the following steps:

15 A. adding a candidate substance to a host cell which harbours the regulatory DNA sequence according to the invention, in particular the 5'-flanking regulatory DNA sequence for the gene for the human catalytic telomerase subunit, or a part region thereof which has a regulatory effect, which sequence or part region is functionally linked to a reporter gene, and

B. measuring the effect of the substance on expression of the reporter gene.

20 The process can be employed for identifying substances which increase the promoter activity, silencer activity or enhancer activity of the catalytic telomerase subunit.

25 The process can furthermore be employed for identifying substances which inhibit the promoter activity, silencer activity or enhancer activator of the catalytic telomerase subunit.

30 The invention furthermore relates to a process for identifying factors which bind specifically to fragments of the DNA fragments according to the invention, in particular the 5'-flanking regulatory DNA sequence of the catalytic telomerase subunit. This method comprises screening an expression cDNA library using the above-described DNA sequence, or subfragments of widely differing length, as the probe.

The above-described constructs or vectors can also be used for preparing transgenic animals.

5 The invention furthermore relates to a process for detecting telomerase-associated conditions in a patient, which process comprises the following steps:

10 A. incubating a construct or vector, which contains the DNA sequence according to the invention, in particular the 5'-flanking regulatory DNA sequence for the gene for the human catalytic telomerase subunit, or a part region thereof having a regulatory effect, and a reporter gene, with body fluids or cell samples,

15 B. detecting the activity of the reporter gene in order to obtain a diagnostic value; and

20 C. comparing the diagnostic value with standard values for the reporter gene construct in standardized normal cells or body fluids of the same type as the test sample;

The detection of diagnostic values which are higher or lower than the standard comparative values indicates a telomerase-associated condition, which in turn indicates a pathogenic condition.

25 Explanation of the figures:

Fig. 1: Southern blot analysis using genomic DNA from various species

30 A: Photograph of an ethidium bromide-stained 0.7% agarose gel containing approximately 4 µg of Eco RI-cut genomic DNA. Track 1 contains Hind III-cut λ DNA as size markers (23.5, 9.4, 6.7, 4.4, 2.3, 2.0 and 0.6 kb). Tracks 2 to 10 contain human, rhesus monkey, Sprague

Dawley rat, BALB/c mouse, dog, bovine, rabbit, chicken and yeast (*Saccharomyces cerevisiae*) genomic DNA.

B: Autoradiogram, corresponding to Fig.1 A, of a Southern blot analysis in which radioactively labelled hTC-cDNA probe of about 720 bp in length is used for the hybridization.

Fig. 2: Restriction analysis of the recombinant  $\lambda$  DNA of the phage clone P12, which hybridizes with a probe from the 5' region of the hTC cDNA.

The figure shows a photograph of an ethidium bromide-stained 0.4% agarose gel. Tracks 1 and 2 contain Eco RI/Hind III-cut  $\lambda$  DNA and a 1 kb ladder from Gibco as size markers. Tracks 3 - 7 each contain 250 ng of the DNA from the recombinant phage which has been cut with Bam HI (track 3), Eco RI (track 4), Sal I (track 5), Xho I (track 6) and Sac I (track 7). The arrows mark the two  $\lambda$  arms of the vector EMBL3 Sp6/T7.

Fig. 3: Restriction analysis and Southern blot analysis of the recombinant  $\lambda$  DNA of the phage clone which hybridizes with a probe from the 5' region of the hTC cDNA.

A: The figure shows a photograph of an ethidium bromide-stained 0.8% agarose gel. Tracks 1 and 15 contain a 1 kb ladder from Gibco as size markers. Tracks 2 to 14 each contain 250 ng of cut  $\lambda$  DNA from the recombinant phage clone. The following enzymes were employed: track 2: Sac I, track 3: Xho I, track 4: Xho I, Xba I, track 5: Sac I, Xho I, track 6: Sal I, Xho I, Xba I, track 7: Sac I, Xho I, Xba I, track 8: Sac I, Sal I, Xba I, track 9: Sac I, Sal I, BamH I, track 10: Sac I, Sal I, Xho I, track 11: Not I, track 12: Sma I, track 13: empty, track 14: not digested.

B: Autoradiogram, corresponding to Fig. 3 A, of a Southern blot analysis. A 5'-hTC cDNA fragment of about 420 bp in length was used as the probe for the hybridization.

5      Fig. 4:      Partial DNA sequence of the 5'-flanking region and of the promoter of the gene for the human catalytic telomerase subunit. The ATG start codon in the sequence is printed in bold. The depicted sequence corresponds to SEQ ID NO 1.

10      Fig. 5:      Use of primer extension analysis to identify the transcription start.

The figure shows an autoradiogram of a denaturing polyacrylamide gel which was selected for depicting a primer extension analysis. An oligonucleotide                      having                      the                      sequence

15      5'GTTAAGTTGTAGCTTACACTGGTTCTC 3' was used as the primer. The primer extension reaction was loaded in track 1. Tracks G, A, T and C constitute the sequence reactions using the same primer and the corresponding dideoxynucleotides. The thick arrow marks the main transcription start while the thin arrows point to three subsidiary transcription start points.

20

Fig. 6:      cDNA sequence of the human catalytic telomerase subunit (hTC; cf. our pending application PCT/EP/98/03468). The depicted sequence corresponds to SEQ ID NO 2.

25

Fig. 7:      Structural organization and restriction map of the human hTC gene and its 5'-flanking and 3'-flanking regions.

30      Exons are shown as consecutively numbered rectangles which are filled-in in black, and introns are shown as regions which are not filled in. Untranslated sequence segments in the exons are hatched. Translation starts in exon 1 and ends in exon 16. Restriction enzyme cleavage sites

are marked as follows: S, SacI; X, XhoI. The relative arrangement of the five phage clones (P2, P3, P5, P12, P17), and of the product from the genome walking, are shown by thin lines. As the dots indicate, the sequence of intron 16 has only been partly deciphered.

5

Fig. 8: HTL splice variants.

A: Diagrammatic structure of the hTC mRNA splice variants. The complete hTC mRNA is depicted as a rectangle with a grey background in the upper region of the figure. The 16 exons are depicted in accordance with their size. The translation start (ATG) and the stop codon, and also the telomerase-specific T motif, and the seven RT motifs, are all shown. The hTC variants are subdivided into deletion and insertion variants. The missing exon sequences are marked in the deletions. The insertions are shown by additional white rectangles. The sizes and origins of the inserted sequences are given. Newly formed stop codons are marked. The size of the insertion in variant INS2 is unknown.

10

15

B: Exon-intron transitions in the hTC splice variants. Unspliced 5'-flanking and 3'-flanking sequences are shown as white rectangles. The origins of the exon and intron sequences are given. Intron and exon sequences are shown in small letters and large letters, respectively. The donor and acceptor sequences in the splice sites are underlaid as grey rectangles, and their exon and intron origins are also given.

20

25

Fig. 9: Identification of the transcription start by means of RT-PCR analysis.

The RT-PCR was carried out using a cDNA library prepared from HL 60 cells and genomic DNA as the positive control. A common 3' primer hybridizes to a region of the exon 1 sequence. The positions of the different 5' primers in the coding region or the 5'-flanking region are given. In the negative control, no template DNA was added to the PCR reaction. M: DNA size marker.

30

Fig. 10: Nucleotide sequence and structural features of the hTC promoter.

The figure depicts 11273 bp of the 5'-flanking hTC gene sequence, beginning with the translation start codon ATG (+1). The putative region of the translation start is underlined. Possible regulatory sequence segments within the 4000 bp upstream of the translation start are ringed. The depicted sequence corresponds to SEQ ID NO 3.

Fig. 11: Activity of the hTC promoter in HEK-293 cells.

The first 5000 bp of the 5'-flanking hTC gene region are shown diagrammatically in the upper part of the figure. The ATG start codon is picked out. CpG-rich islands are marked by grey rectangles. The sizes of the hTC promoter-luciferase construct are shown on the left-hand side of the figure. The promoterless pGL2 basic construct and the SV40 promoter construct pGL2-Pro were used as controls in each transfection. The relative luciferase activities of the different promoter constructs in HEK cells are shown as continuous bars on the right-hand side of the figure. The standard deviation is indicated. The numerical values represent the average of two independent experiments which were carried out in duplicate.

Tab. 1: Exon-intron transitions in the hTC gene

The table lists the nucleotide sequences at the 3' and 5' splice transitions of the hTC gene. The consensus sequences for donor and acceptor sequences (AG and GT) are underlaid with grey rectangles. The table shows the intron sequences (small letters) and exon sequences (large letters) which flank the splice acceptor and donor sites. The sizes of the exons and introns are given in bp.

Tab. 2: Potential binding sites for DNA-binding factors in the nucleotide sequence of intron 2



The search for possible DNA-binding factors (e.g. transcription factors) was carried out using the "find pattern" algorithm from the Genetics Computer Group (Madison, USA) GCG sequence analysis program package. The table lists the abbreviations of the DNA-binding factors which were identified and their location in intron 2.

Tab. 1

3' Acceptor Sequence			5' Donor Sequence		
Intron	Exon	Exon bp	Intron	Exon	Intron bp
		No.			No.
5' flanking region	GTTCAGGCAGCGTGGT	1	CGCCCTCTCTTCGCCAG	281	1
caggcgcttcccccgag	GTGTCCTGCCCTGAAGAGC	2	TGGCTGGCAGGAGCCCCAG	1354	2
catgtccttctcgtttaag	GGGTGGCTGTGTTCCGGC	3	TGCAAGCATTTGGAATCAG	196	3
gagggctctctattgag	ACAGCACTTGAAGAGGGTG	4	GTTCGCAGAGAAAAGAGG	181	4
cccatgtgtcccccgag	GCCGAGCGTCTCACCTCGA	5	TGAGCTGTACTTTGTCAAG	180	5
ctgcctccactcacagag	GTGATGTGACGGGGCGGT	6	CAAGGCTTCAAGAGCCAC	156	6
ccctctctctgcgggag	GTCTTACCTTGACAGACC	7	TGCCGTGTCATCGAGCAG	96	7
ctcccgctctgtttcgag	AGCTCCTCCCTGAATGAGG	8	CCGTGGCATCAGGGGCAA	86	8
ctgtgtcttcccccgag	GTCTACGTCCAGTGCCAG	9	CGGGGATTCGGCGGGACGG	114	9
gtattttcccttatttag	GCTGCTCCTGCGTTTGGTG	10	ACGCGAAAACCTTCCTCAG	72	10
cattgccccctctgccttag	GACCTGGTCCGAGGTGTC	11	TGCAGAGCGACTACTCCAG	189	11
attccccctgtgtctdag	CTATGCCCCGACCTCCATC	12	CCTGTTTCTGGATTTCAG	127	12
tcctttctggcgactctag	GTGAACAGCCCTCCAGACGG	13	TCCTGCTGCAGGCGTACAG	62	13
ctgtccgcacatcctctdag	GTTTCAGGCATGTGTGCTG	14	CTGAAGCCCAAGAACGCAG	125	14
agcctctgttttccccdag	GGATGTCGCTGGGGGCCAA	15	CTGGGGTCACTCAGGACAG	138	15
tctgattttggcccccgag	CCCAGAGCAGCTGAGTCG	16	TTTTTTCAGTTTGAATAAA	664	16
			3' flanking region		
			gtgggcctccccgggtgag		104
			gtgaggaggtgtggccgt		8616
			gtactgtatccccacgcca		2089
			gtggctgtgctttgttta		687
			gtgggtgccggggaccccc		494
			gttaaggttcaagtgtgata		>4660
			gtttgggcaactgccctgca		980
			gtgagtcaggtggccaggt		2485
			gtgaggcctcctctctcccc		1984
			gtgagggcccggtgccgtgtg		1871
			gtgagcgcacctggccgga		3801
			gtgagcaggctgatgtca		880
			gtgagccgccaccaagggg		3187
			gtatgtgcaggtgcctggc		781
			gcaagtggtgggtggaggcc		536

Tab. 2

Factors	Location in intron 2
C/EBP	2925
CRE.2	2749
Sp1	2378, 4094, 4526, 4787, 4835, 4995
AP-2 CS3	5099
AP-2 CS4	2213, 3699, 4667, 5878, 5938, 6059, 6180, 6496
AP-2 CS5	5350, 5798, 5880, 5940, 6061, 6182, 6375, 6498
PEA3	934, 2505
P53	2125
GR uteroglobin	848, 1487, 2956
PR uteroglobin	3331
Zeste-white	1577, 1619, 1703, 1745, 1787, 1829, 1871, 1913, 1955, 1997, 2039, 2081, 3518, 3709, 4765, 5014, 5055
GRE	846
MyoD-MCK right site/rev	447, 509, 558, 1370, 1595, 1900, 2028, 2099, 4557
MyoD-MCK left site	108, 118, 453, 1566, 1608, 1692, 1734, 1818, 1902, 1986, 2372, 2460, 2720, 3491, 5030
Ets-1 CS	6408
API	3784, 4406
CREB	2801
GATA-1	839, 1390, 3154
c-Myc	108, 118, 453, 1566, 1608, 1692, 1734, 1818, 1902, 1986, 2372, 2460, 2720, 3491, 5030
CACCC site	991
CCAAT site	1224
CCAC box	992
CAAT site	463, 2395
Rb site	992, 4663
TATA	3650
CDEI	106, 1564, 1606, 1690, 1732, 1816, 1900, 1984

### Examples

The human gene for the catalytic telomerase subunit (ghTC), and the regions of this gene located 5' and 3', were cloned, while the start point for transcription was determined, potential binding sites for DNA-binding proteins were identified and active promoter fragments were highlighted. The sequence of the hTC cDNA (Fig. 6) has already been reported in our application PCT/EP/98/03468, which is also pending. Unless otherwise mentioned, all the data refer to the position of the cDNA in this sequence.

#### Example 1

A genomic Southern blot analysis was used to determine whether ghTC constitutes a single gene in the human genome or whether there exist several loci for the hTC gene and possibly also ghTC pseudogenes.

In order to do this, a commercially available zoo blot from Clontech was subjected to Southern blot analysis. This blot contains 4 µg of Eco RI-cut genomic DNA from nine different species (human, monkey, rat, mouse, dog, bovine, rabbit, chicken and yeast). With the exception of yeast, chicken and human, the DNA was isolated from kidney tissue. The human genomic DNA was isolated from placenta and the chicken genomic DNA was purified from liver tissue. An hTC cDNA fragment of about 720 bp in length, which was isolated from hTC cDNA, variant Del2 (position 1685 to 2349 plus 2531 to 2590 in Fig. 6 [deletion 2; cf. Example 5 in Fig. 8]), was used as the radioactively labelled probe in the autoradiogram in Fig. 1. The experimental conditions for the blot hybridization and washing steps were taken from Ausubel *et al.* (1987).

In the case of the human DNA, the probe recognizes two specific DNA fragments. The smaller Eco RI fragment, of from about 1.5 to 1.8 kb in length, probably originates from two Eco RI cleavage sites in an intron in the ghTC DNA. On the

basis of this result, it is to be assumed that only one single ghTC gene is present in the human genome.

### Example 2

5

In order to isolate the 5' flanking hTC gene sequence, approx.  $1.5 \times 10^6$  phages from a human genomic placenta gene library (EMBL 3 SP6/T7 from Clontech, order number HL1067j) were hybridized on nitrocellulose filters (0.45  $\mu\text{m}$ ; from Schleicher and Schuell), in accordance with the manufacturer's instructions, with a radioactively labelled 5'-hTC cDNA fragment of about 500 bp in length (position 839 to 1345 in Fig. 6). The nitrocellulose filters were firstly incubated, at 42°C for two hours, in 2 x SSC (0.3 M NaCl; 0.5 M Tris-HCl, pH 8.0) and then in a prehybridization solution (50% formamide; 5 x SSPE, pH 7.4; 5 x Denhard's solution; 0.25% SDS; 100  $\mu\text{g}$  of herring sperm DNA/ml). For the overnight hybridization, the prehybridization solution was supplemented with  $1.5 \times 10^6$  cpm of denatured, radioactively labelled probe/ml of solution. Nonspecifically bound radioactive DNA was removed under stringent conditions, i.e. by means of three five-minute steps of washing with 2 x SSC; 0.1% SDS at from 55 to 65°C. The filters were evaluated by autoradiography.

20

The phage clones which were identified in this primary investigation were purified (Ausubel *et al.* (1987)). In subsequent analyses, one phage clone, i.e. P12 turned out to be potentially positive. A  $\lambda$  DNA preparation carried out on this phage (Ausubel *et al.* (1987)), and the subsequent restriction digestion with enzymes which release the genomic insert in fragments, showed that this phage clone contains an insert of approx. 15 kb in the vector (Fig. 2).

25

In order to isolate the complete hTC gene sequence, in each case from 1 to  $1.5 \times 10^6$  phages were screened, in independent experiments, with in each case different radioactively labelled probes, as described above.

30

The phage clones which were identified in these primary investigations, and which were positive for the corresponding probes, were purified. The phage clone P17 was found to contain an hTC cDNA fragment of about 250 bp in length (position 1787 to 2040 in Fig. 6). The phage clone P2 was identified as containing an hTC cDNA  
5 fragment of about 740 bp in length (position 1685 to 2349 plus 2531 to 2607 in Fig. 6 [deletion 2; cf. Example 5]). The phage clones P3 and P5 were found to contain a 3' hTC cDNA fragment of 420 bp in length (position 3047 to 3470 in Fig. 6). After the  $\lambda$  DNA had been prepared from these phages, and subsequently subjected to restriction digestion with enzymes which release the genomic insert in fragments, the  
10 inserts were subcloned into plasmids (Example 4).

### **Example 3**

In order to investigate whether the 5' end of the hTC cDNA was also present in the  
15 insert in the recombinant phage clone P12, the  $\lambda$  DNA from this clone was hybridized, in a Southern blot analysis, with a radiactively labelled hTC cDNA fragment of about 440 bp in length (position 1 to 440 in Fig. 6) from the extreme 5' region (Fig. 3).

20 Since the isolated  $\lambda$  DNA from the positive clone also hybridizes with the extreme 5' end of the hTC cDNA, this phage probably also contains the 5' sequence region flanking the ATG start codon.

### **Example 4**

25 In order to subclone the entire 15 kb insert in the positive phage clone P12 in the form of subfragments, and subsequently to sequence these fragments, restriction endonucleases which, on the one hand, release the entire insert from EMBL3 Sp6/T7 (cf. Example 2) and, in addition, cut within the insert, were selected for digesting the  
30 DNA.

In all, two Xho I subfragments, of about 8.3 and about 6.5 kb in length, respectively, and three Sac I subfragments, of about 8.5, about 3.5 and about 3 kb in length, respectively, were subcloned into the pBluescript KS(+) vector (from Stratagene). The 5123 bp 5'-flanking nucleotide sequence of the ghTC gene region, starting from the ATG start codon, was determined by analysing the sequences of these fragments (Fig. 4; corresponding to SEQ ID NO 1). Fig. 4 depicts the first 5123 bp (starting from the ATG start codon). Fig. 10 depicts the entire cloned 5' sequence (corresponding to SEQ ID NO 3).

In order to subclone the entire insert, of approx. 14.6 kb in size, in phage clone P17 in the form of subfragments, restriction endonucleases which, on the one hand, release the entire insert from EMLB3 Sp6/T7 and, in addition, cut a few times within the insert, were selected for digesting the DNA. Three XhoI/BamHI fragments, of 7.1 kb, 4.2 kb and 1.5 kb in size, respectively, and one BamHI fragment, of 1.8 kb in size, were subcloned by means of using a combination digestion with the enzymes XhoI and BamHI. Combination restriction digestion with the enzymes XhoI and XbaI resulted in a XhoI/XbaI fragment of 6.5 kb in size, and two XhoI fragments, of 6.5 kb and 1.5 kb in size, respectively, being cloned.

Digestion with the restriction enzyme XhoI was used to subclone the insert, of approx. 17.9 kb in size, in phage clone P2 in the form of subfragments. In all, three XhoI subfragments, of 7.5 kb, 6.4 kb and 1.6 kb in length, respectively, were cloned. Four SacI fragments, of 4.8 kb, 3 kb, 2 kb and 1.8 kb in size, respectively, were additionally subcloned by digesting with the restriction enzyme SacI.

The insert, of approx. 13.5 kb in size, in phage clone P3 was subcloned by digesting with the restriction enzymes SacI and/or XhoI. Six SacI subfragments, of 3.2 kb, 2 kb, 0.9 kb, 0.8 kb, 0.65 kb and 0.5 kb in length, respectively, and two XhoI subfragments, of 6.5 kb and 4.3 kb in length, respectively, were obtained in this connection.

The insert, of approx. 13.2 kb in size, in phage clone P5 was subcloned by digesting with the restriction enzymes *SacI* and/or *XhoI*. In all, *SacI* fragments of 6.5 kb, 3.3 kb, 3.2 kb, 0.8 kb and 0.3 kb in size, and *XhoI* fragments of 7 kb and 3.2 kb in size, were subcloned.

5

In order to clone the hTC genomic sequence region located 3' of phage clone P17 and 5' of phage clone P2, 3 genomic walkings were carried out using the Clontech GenomeWalker™ kits (catalogue number K1803-1) and various combinations of primers. In a final volume of 50 µl, 10 pmol of dNTP mix were added to 1 µl of human GenomeWalker Library HDL (from Clontech), and a PCR reaction was carried out in 1xKlen Taq PCR reaction buffer and 1xAdvantage Klen Taq polymerase mix (from Clontech). 10 pmol of an internal gene-specific primer, and 10 pmol of the adaptor primer AP1 (5'-GTAATACGACTCACTATAGGGC-3'; from Clontech) were added as primers. The PCR was carried out in 3 steps as a touchdown PCR. First of all, denaturation was carried out at 94°C for 20 sec, and the primers were then annealed, and the DNA chain extended, at 72°C for 4 min, over 7 cycles. There then followed 37 cycles in which the DNA was denaturated at 94°C for 20 sec but the subsequent primer extension took place at 67°C for 4 min. In conclusion, there followed a chain extension at 67°C for 4 min. After this first PCR, the PCR product was diluted 1:50. One µl of this dilution was used in a second nested PCR together with 10 pmol of dNTP mix in 1xKlen Taq PCR reaction buffer and 1xAdvantage Klen Taq polymerase mix and also 10 pmol of a nested gene-specific primer and 10 pmol of the nested Marathon Adaptor primers AP2 (5'-ACTATAGGGCACGCGTGGT-3'; from Clontech). The PCR conditions corresponded to the parameters which were selected in the first PCR. As the sole exception, only 5 cycles rather than 7 cycles were selected in the first PCR step and only 24 cycles, instead of 37 cycles, were run in the second PCR step. The products of this nested genomic walking PCR were cloned into the TA Cloning Vector pCR11 from InVitrogen.

30



In the first genomic walking, the gene-specific primer C3K2-GSP1 (5'-GACGTGGCTCTTGAAGGCCTTG-3') and the nested gene-specific primer C3K2-GSP2 (5'-GCCTTCTGGACCACGGCATAACC-3') were used, together with the HDL library 4, and a PCR fragment of 1639 bp in length was obtained. In the second genomic walking, a PCR fragment of 685 bp in length was amplified from the HDL library 4 using the gene-specific primer C3F2 (5'-CGTAGTTGAGCACGCTGAACAGTG-3') and the nested gene-specific primer C3F (5'-CCTTCACCCTCGAGGTGAGACGCT-3. The third genomic walking mixture, using the gene-specific primer DEL5-GSP1 (5'-GGTGGATGTGACGGGCGCGTACG-3') and the nested gene-specific primer C5K-GSP1 (5'-GGTATGCCGTGGTCCAGAAGGC-3'), led to a 924 bp PCR fragments being cloned from the HDL library 1. In all, 2100 bp of the genomic hTC region located 3' of phage clone P17 were identified using this genomic walking method (see Fig. 7).

The subcloned fragments, and the genomic walking products, were sequenced in single-stranded form. The Lasergene Biocomputing Software (DNASTAR Inc. Madison, Wisconsin, USA) was used to identify overlapping regions and form contigs. In all, 2 large contigs were assembled from the sequences collected from phage clones P12, P17, P2, P3 and P5, and also the sequence data from the genomic walking. Contig 1 consists of sequence data from phage clones P12 and P17 and the sequence data from the genomic walking. Contig 2 was put together from the sequences from phage clones P2, P3 and P5. Overlapping phage clone regions are shown diagrammatically in Fig. 7. The sequence data from the 2 contigs are shown below. The ATG start codon in contig 1 is underlined. The TGA stop codon is underlined in contig 2.

## Contig1:

	ACTTGAGCCC	AAGAGTTCAA	GGCTACGGTG	AGCCATGATT	GCAACACCAC	ACGCCAGCCT	TGGTGACAGA	70
5	ATGAGACCCT	GTCTCAAAAA	AAAAAAAAAA	AATTGAAATA	ATATAAAGCA	TCTTCTCTGG	CCACAGTGGA	140
	ACAAAACCCAG	AAATCAACAA	CAAGAGGAAT	TTTGAAAACT	ATACAAACAC	ATGAAAATTA	AACAATATAC	210
	TTCTGAATGA	CCAGTGAGTC	AATGAAGAAA	TTAAAAAGGA	AATTGAAAAA	TTTATTTAAG	CAAAATGATA	280
	CGGAAACATA	ACCTCTCAAA	ACCCACGGTA	TACAGCAAAA	GCAGTGCTAA	GAAGGAAGTT	TATAGCTATA	350
	AGCAGATACA	TCAAAAAAGT	AGAAAAGCCA	GGCCGAGTGG	CTCATGCCTG	TAATCCCAGC	ACTTTGGGAG	420
10	GCCAAGGCGG	GCAGATCGCC	TGAGGTACAG	AGTTTCGAGC	CAGCCTGACC	AACACAGAGA	AACCTTGTCG	490
	CTACTAAAAA	TACAAAATTA	GCTGGGCATG	GTGGCACATG	CCTGTAATCC	CAGTACTCTG	GGAGGCTGAG	560
	GCAGGATAAC	CGCTTGAACC	CAGGAGGTGG	AGGTTGCGGT	GAGCCGGGAT	TGCCCCATTG	GACTCCAGCC	630
	TGGGTAAACA	GAGTGAAACC	CTGTCTCAAG	AAAAAAAAAA	AAGTAGAAAA	ACTTAAAAAT	ACAACCTAAT	700
	GATGCACCTT	AAAGAACTAG	AAAAGCAAGA	GCAAACATAA	CCTAAAATTG	GTAAAAGAAA	AGAAAATAA	770
15	AAAGATCAGAG	CAGAAATATA	TGAAACTGAA	AGATAACAAT	ACAAAAGATC	AACAAAATTA	AAAGTTGGTT	840
	TTTTGAAAAG	ATAAACAAAA	TTGACAAACC	TTTGCCACAG	CTAAGAAAAA	AGGAAAGAAG	ACCTAAATAA	910
	ATAAAGTCAG	AGATGAAAAA	AGAGACATTA	CAACTGATAC	CACAGAAATT	CAAAGGATCA	CTAGAGGCTA	980
	CTATGAGCAA	CTGTACACTA	ATAAATTGAA	AAACCTAGAA	AAAATAGATA	AATTCCTAGA	TGCATACAAC	1050
	CTACCAAGAT	TGAACCATGA	AGAAATCCAA	AGCCCAACAA	GACCAATAAC	AATAATGGGA	TTAAAGCCAT	1120
20	AATAAAAAAGT	CTCCTAGCAA	AGAGAAGCCC	AGGACCCAA	GGCTTCCCTG	CTGGATTTTA	CCAATCATTT	1190
	AAAGAAGAAT	GAATTCCAAT	CCTACTCAAA	CTATTCTGAA	AAATAGAGGA	AAGAATACTT	CCAAACTCAT	1260
	TCTACATGCG	CAGTATTACC	CTGATTCCAA	AACCAGACAA	AAACACATCA	AAAAACAAAC	AACAAAAAAA	1330
	CAGAAAGAAA	GAAAACTACA	GGCCAATATC	CCTGATGAAT	ACTGATACAA	AAATCCTCAA	CAAAACACTA	1400
	GCAAAACCAA	TTAAACAACA	CCTTCGAAG	AGATTCAATT	GTGATCAAGT	GGGATTTATT	CCAGGGATGG	1470
25	AAGGATGGTT	CAACATATGC	AAATCAATCA	ATGTGATACA	TCATCCCAAC	AAAAATGAAGT	ACAAAAACTA	1540
	TATGATTATT	TCACCTTATG	CAGAAAAAGC	ATTTGATAAA	ATTCTGCACC	CTTCATGATA	AAAACCTCTA	1610
	AAAAACCCAG	TATACAAGAA	ACATACAGGC	CAGGCACAGT	GGCTCACACC	TGCGATCCCA	GCCTCTGGG	1680
	AGGCCAAGGT	GGGATGATTG	CTTGGGCCCA	GGAGTTTGAG	ACTAGCCTGG	GCAACAAAAT	GAGACCTGGT	1750
	CTACAAAAAA	CTTTTTTAAA	AAATTAGCCA	GGCATGATGG	CATATGCCTG	TAGTCCCAGC	TAGTCTGGAG	1820
30	GCTGAGGTGG	GAGAATCAGT	TAAGCCTAGG	AGGTCGAGGC	TGCAGTGAGC	CATGAACATG	TCAGTCTGCT	1890
	CCAGCCTAGA	CAACAGAACA	AGACCCCACT	GAATAAGAAG	AAGGAGAAGG	AGAAGGGAGA	AGGGAGGGAG	1960
	AAGGGAGGAG	GAGGAGAAGG	AGGAGGTGGA	GGAGAAGTGG	AAGGGGAAGG	GGAAGGGAAA	GAGGAAGAAG	2030
	AAGAAACATA	TTTCAACATA	ATAAAAGCCC	TATATGACAG	ACCGAGGTAG	TATTTATAGG	AAAAACTGAA	2100
	AGCCTTTCTCT	CTAAGATCTG	GAAATGACA	AGGGCCCACT	TTCAACCACTG	TGATTCACAA	TAGTACTAGA	2170
35	AGTCCTAGCT	AGAGCAATCA	GATAAGAGAA	AGAAATAAAA	GGCATCCAAA	CTGGAAAGGA	AGAAGTCAAA	2240
	TTATCTGTGT	TGAGATGAT	ATGACTTTAT	ATCTGGAAAA	GACTTAAGAC	ACCACTAAAA	ACTATTATGA	2310
	GCTGAAATTT	GGTACAGCAG	GATACAAAAT	CAATGTACAA	AAATCAGTAG	TATTTCTATA	TTCCAACAGC	2380
	AAACAATCTG	AAAAAGAAAC	CAAAAAAGCA	GCTACAAATA	AAATTAACA	GCTAGGAATT	AACCAAGAA	2450
	GTGAAGAGAT	TCTACAATGA	AAACTATAAA	ATGTTGATAA	AAGAAATTGA	AGAGGGCACT	AGAAAAGAAA	2520
40	AGATATTCCA	TGTTCATAGA	TTGGAAGAAT	AAATACTGTT	AAAATGTCCA	TACTACCCAA	AGCAATTTAC	2590
	AAATTCAATG	CAATCCCTAT	TAAATACTA	ATGACGTTCT	TCACAGAAAT	AGAAGAAACA	ATTCTAAGAT	2660
	TTGTACAGAA	CCACAAAAGA	CCAGAATAG	CCTGACCAAA	AAGAACAAAA	CTGGAGGCAT	TGGAAGCAT	2730
	CACATTACCT	GACTTCAAT	TATACTACAA	AGCTATAGTA	ACCCAAACTA	CATGGTACTG	GCATAAAAA	2800
	AGATGAGACA	TGGACCAGAG	GAACAGAATA	GAGAATCCAG	AAACAAATCC	ATGCATCTAC	AGTGAACCTA	2870
45	TTTTTGACAA	AGGTGCCAAG	AACATCTTTT	GGGGAAAAAG	TAATCTCTTC	AATAAATGGT	GCTGGAGGAA	2940
	CTGGATATCC	ATATGCAAAA	TAACAATACT	AGAACTCTGT	CTCTCACCAT	ATACAAAAGC	AAATCAAAAT	3010
	GGATGAAAGG	CTTAAATCTA	AAACCTCAAA	CTTTGCAACT	ACTAAAAGAA	AACACCCGGG	AAACTCTCCA	3080
	GGACATTGGA	GTGGGCAAG	ACTTCTTGAG	TAATTCCTCTG	CAGGCACAGG	CAACCAAGGC	AAAAACAGAC	3150
	AAATGGGATC	ATATCAAGTT	AAAAAGCTTC	TGCCAGGCAA	AGGAAACAAT	CAACAAAGAG	AAGAGACAAC	3220
50	CCACAGAATG	GGAGAATATA	TTTGCAAACT	ATTCATCTAA	CAAGGAATTA	ATAACCACTA	TATATAAGGA	3290
	GCTCAAACTA	CTCTATAAGA	AAAACACCTA	ATAAGCTGAT	TTTCAAAAAT	AAGCAAAAAG	CTGGGTGAGA	3360
	CATTTCTCAA	AAATAAGTCAT	ACAAATGGCA	AACAGGCATC	TGAAAATGTG	CTCAACACCA	CTGATCATCA	3430
	GAGAAATGCA	AATCAAAACT	ACTATGAGAG	ATCATCTCAT	CCCAGTTAAA	ATGGCTTTTA	TTCAAAAGAC	3500
	AGGCAATTAAC	AAATGCCAGT	GAGGATGTGG	ATAAAAGGAA	ACCCTTGGAC	ACTGTTGGTG	GAAATGGAAA	3570
55	TTGCTACCAC	TATGGAGAAC	AGTTTGAAAG	TTCTCAAAA	AACTAAAAAT	AAAGCTACCA	TACAGCAATC	3640
	CCATTGCTAG	GTATATACTC	CAAAAAAGGG	AATCAGTGTA	TCAACAAGCT	ATCTCCACTC	CCACATTTAC	3710
	TGCAGCACTG	TTCATAGCAG	CCAAGGTTTG	GAAAGCAACT	CAGTGTCAT	CAACAGACGA	ATGGAAGGAA	3780
	AAAATGTGGT	GCACATACAC	AATGGAGTAC	TACGCAGCCA	TAAAAAAGAA	TGAGATCCTG	TCAGTTGCAA	3850
	CACCATGGGG	GGCACTGGTC	AGTATGTTAA	GTGAAATAAG	CCAGGCACAG	AAAGACAAAC	TTTTCATGTT	3920
60	CTCCCTTACT	TGTGGGAGCA	AAAATTAAAA	CAATTGACAT	AGAAATAGAG	GAGAATGGTG	GTCTAGAGG	3990
	GGTGGGGGAG	AGGGTGACTA	GAGTCAACAA	TAATTTATTG	TATGTTTTAA	AATAACTAAA	AGAGTATAAT	4060
	TGGGTTGTTT	GTAACACAAA	GAAAGGATAA	ATGCTTTGAG	GTGACAGATA	CCCCATTTAC	CCTGATGTGA	4130
	TTATTACACA	TTGTATGCCT	GTATCAAAAT	ATCTCATGTA	TGCTATAGAT	ATAAACCTTA	CTATATTAGA	4200
	AATTAATAAT	TTAATGGCCA	GGCACGGTGG	CTCATGTCCG	TAATCCCAGC	ACTTTGGGAG	GCCGAGGGCG	4270
65	GTGGATCACC	TGAGGTGAGG	AGTTTGAAAC	CAGTCTGGCC	ACCATGATGA	AACCCTGTCT	CTACTAAAGA	4340
	TACAAAATTT	AGCCAGGCGT	GGTGGCACAT	ACCTGTAGTC	CCAACACTCA	AGGAGGCTGA	GACAGGAGGA	4410
	TGCTTTGAAC	CTGGGAGGCG	GAGGTTGCAG	TGAGCCGAGA	TCATGCCACT	GCACTGCAGC	CTGGGTGACA	4480
	GAGCAAGACT	CCATCTCAAA	ACAAAAACAA	AAAAAAGAG	ATTAATAATG	TAATTTTTAT	GTACCGTATA	4550
70	AATATATACT	CTACTATATT	AGAAGTTAAA	AATTAAACAA	ATTATAAAG	GTAATTAACC	ACTTAATCTA	4620
	AAATAAGAAC	AATGTATGTG	GGGTTTCTAG	CTTCTGAAGA	AGTAAAGTT	ATGGCCACGA	TGGCAGAAAT	4690
	GTGAGGAGGG	AACAGTGGAA	GTTACTGTTG	TTAGACGCTC	ATACTCTCTG	TAAGTGACTT	AATTTTAAAC	4760
	AAAGACAGGC	TGGGAGAAGT	TAAAGAGGCA	TTCTATAAGC	CCTAAAACAA	CTGCTAAATA	TGGTGAAAGG	4830
	TAATCTCTAT	TAATTACCAA	TAATTACAGA	TATCTCTAAA	ATCGAGCTGC	AGAATTGGCA	CGTCTGATCA	4900
	CACCGCTCCT	TCATTACAGG	TGCTTTTTTT	CTTGTGTGCT	TGGAGATTTT	CGATTGTGTG	TTGCTGTTTG	4970
75	GTTAAACTTA	ATCTGTATGA	ATCCTGAAAC	GAAAAATGGT	GGTGATTTC	TCCAGAGAAA	TTAGAGTACC	5040
	TGGCAGGAAG	CAGGTGGCTC	TGTGGACCTG	AGCCACTTCA	ATCTTCAAGG	GTCTCTGGCC	AAGACCCAGG	5110

	TGCAAGGCAG	AGGCCTGATG	ACCCGAGGAC	AGGAAAGCTC	GGATGGGAAG	GGGCGATGAG	AAGCCTGCCT	5180
	CGTTGGTGAG	CAGCGCATGA	AGTGCCCTTA	TTTACGCTTT	GCAAAGATTG	CTCTGGATAC	CATCTGAAAA	5250
	AGGCGGCCAG	CGGGAATGCA	AGGAGTCAGA	AGCCTCCTGC	TCAAACCCAG	GCCAGCAGCT	ATGGCGCCCA	5320
5	CCCGGGCGTG	TGCCAGAGGG	AGAGGAGTCA	AGGCACCTCG	AAGTATGGCT	TAAATCTTTT	TTTCACCTGA	5390
	AGCAGTGACC	AAGGTGTATT	CTGAGGGGAA	CTTGAGTTAG	GTGCTTCTTT	TAAAAACAGAA	AGTCATGGAA	5460
	GCACCTTCT	CAAGGGAATA	CCAGACGCC	GCTCTCGGCT	CATTTACCTC	TTTCCTCTCT	CCCTCTCTTG	5530
	CCCTCGCGGT	TTCTGATCGG	GACAGAGTGA	CCCCCGTGGA	GCTTCTCCGA	GCCCCGTGCTG	AGGACCCTCT	5600
	TGCAAAGGGC	TCCACAGACC	CCCCGCCCTGG	AGAGAGGAGT	CTGAGCCTGG	CTTAATAACA	AACCTGGGATG	5670
10	TGGCTGGGGG	CGGACAGCGA	CGGCGGGATT	CAAAGACTTA	ATTCCATGAG	TAAATTCAC	CTTCCACAT	5740
	CCGAATGGAT	TTGGATTTTA	TCTTAATATT	TTCTTAAATT	TCATCAAATA	ACATTCAGGA	CTGCAGAAAT	5810
	CCAAAGGCGT	AAAACAGGAA	CTGAGCTATG	TTTGCCAAAG	TCCAAGGACT	TAATAACCAT	GTTCAGAGGG	5880
	ATTTTTCGCC	CTAAGTACTT	TTTATTGGTT	TTCTAAGGT	GGCTTAGGGT	GCAAGGGAAA	GTACACGAGG	5950
	AGAGGCTTGG	GCGGCAGGGC	TATGAGCAGG	GCAGGGCCAC	CGGGGAGAGA	GTCCCCGGCC	TGGGAGGCTG	6020
	ACAGCAGGAC	CACGTACCGT	CCTCCCTGGG	AGCTGCCACA	TTGGGCAACG	CGAAGGCGCG	CACGCTGCGT	6090
15	GTGACTCAGG	ACCCCATACC	GGCTTCTGGG	GCCCAACCCAC	ACTAACCCAG	GAAGTCACGG	AGCTTGAAC	6160
	CCGTGGAAAC	GAACATGACC	CTTGCTGCTG	TGCTTCCCTG	GGTGGGTCAA	GGGTAATGAA	GTGGTGTGCA	6230
	GGAAATGGCC	ATGTAATAAT	CACGACTCTG	CTGATGGGGA	CCGTTCCTTC	CATCATTATT	CATCTTCACC	6300
	CCCAAGGACT	GAATGATTCC	AGCAACTTCT	TCGGGTGTGA	CAAGCCATGA	CAAAACTCAG	TACAAACACC	6370
20	ACTCTTTTAC	TAGGCCACACA	GAGCACGGSC	CACACCCCTG	ATATATTAAG	AGTCCAGGAG	AGATGAGGCT	6440
	GCTTTTCAGC	ACCAGGCTGG	GGTGACAACA	GCGGCTGAAC	AGTCTGTTC	TCTAGACTAG	TAGACCTGG	6510
	CAGGCACTCC	CCCAGATTCT	AGGGCCTGGT	TGCTGCTTCC	CGAGGGCGCC	ATCTGCCCTG	GACTCTCAGC	6580
	CTGGGGTGCC	ACACTGAGGC	CAGCCCTGTC	TCCACACCTT	CCGCCCTCCAG	GCCTCAGCTT	CTCCAGCAGC	6650
	TTCTTAACAC	CTGGGTGGGC	CGTGTTCCAG	CGCTACTGTC	TCACCTGTCC	CACCTGTGCT	TGTCTCAGCG	6720
25	ACGTAGCTCG	CACGGTTCCT	CCTCACATGG	GGTGTCTGTC	TCCTTCCCCA	ACACTCACAT	CGCTTGAAG	6790
	GAGGAGATTG	TGCGCCTCCC	AGACTGGCTC	CTCTGAGCCT	GAACCTGGCT	CGTGGCCCCC	GATGCAGGTT	6860
	CCTGGCGTCC	GGCTGCACGC	TGACCTCCAT	TTCCAGGCGC	TCCCGTCTC	CTGTCACTGT	CCGGGGCCTG	6930
	CCGGTGTGTT	CTTCTGTGTT	TGTGCTCTTT	TCCAGTCTTC	GCTGCGTGTG	TCTCTGCCCG	GAGGCTCTC	7000
	GGGGTTTTTA	TAGGCATAGG	ACGGGGGCGT	GGTGGGCCAG	GGCGCTCTTG	GGAAATGCAA	CATTTGGGTG	7070
30	TGAAAGTAGG	AGTGCCGTGC	CTCACCTTAG	TCCACGGGCA	CAGGCTGGG	GATGGAGCCC	CCGCCAGGGA	7140
	CCCGCCCTTC	TCTGCCAGC	ACTTCTCTGC	CCCCCTCCCT	CTGGAACACA	GAGTGGCAGT	TCCACAAGC	7210
	ACCTAAGCATC	CTCTTCCCAA	AAGACCCAGC	ATTGGCACCC	CTGGACATTT	GCCCCACAGC	CCTGGGAATT	7280
	CACGTACTGA	CGCACATCAT	GTACACACTC	CCGTCCACGA	CCGACCCCGC	CTGTTTTATT	TTAATAGTGA	7350
	CAAAGCAGGG	AAATCCCTGC	TAAATGTCTC	TTTAACAAAC	TGGTTAAACA	AACGGGTCCA	TCCGCACGGT	7420
35	GGACAGTTCC	TCACAGTGAA	GAGGAACATG	CCGTTTATAA	AGCCTGCAGG	CATCTCAAGG	GAATTACGCT	7490
	GAGTCAAAAC	TGCCACCTCC	ATGGGATACG	TACGCAACAT	GCTCAAAAAG	AAAGAATTTC	ACCCCATGGC	7560
	AAGGGAGTGG	TTAGGGGGGT	TAAGGACGGT	GGGGGGGGCA	GCTGGGGGCT	ACTGCACGCA	CTCTTTACTA	7630
	AAGCCAGTTT	CTTGTTCTGT	ATGGTATTGG	CTCAGTTATG	GGAGACTAAC	CATAGGGGAG	TGGGGATGGG	7700
	GGAAACCGGA	GGCTGTGCCA	TCTTTGCCAT	GCCCCAGTGT	CCTGGGCAGG	ATAATGCTCT	AGATGTCCCG	7770
40	ACGTCTGTAG	TCCCCAAAC	CTGTGGACAG	AACCCGGCCG	GCCCCAGGGC	CTTTGACAGT	GTGATCTCCC	7840
	TGAGGACCCT	GAGGTCTGGG	ATCCTTCCGG	ACTACCTGCA	GGCCCCGAAA	GTAATCCAGG	GGTCTGCGGA	7910
	AGAGGCGGGC	AGGAGGGTCA	GAGGGGGGCA	GCCTCAGGAC	GATGGAGGCA	GTCAGTCTGA	GGGTGAAAAG	7980
	GGAGGGAGGG	CCTCGAGCCC	AGGCCTGCAA	GCGCCTCCAG	AAGCTGGAAA	AAGCGGGGAA	GAGGCTCTCC	8050
	ACGGAGCCTG	CAGCAGGAAG	GCACGGCTGG	CCCTTAGCCC	ACCAGGGCCC	ATCGTGGACC	TCCGGCCTCC	8120
	GTGCCATAGG	AGGGCACTCG	CGCTGCCCTT	CTAGCATGAA	GTGTGTGGGG	ATTTGCAGAA	GCAACAGGAA	8190
45	ACCCATGCAC	TGTGAATCTA	GGATTATTTC	AAAACAAGAG	TTTACAGAAA	CATCCAAGGA	CAGGGCTGCA	8260
	GTGCCCTCCG	GCAAGGGCAG	GGCAGGCACG	AGTGATTTTA	TTTAGCTATT	TTATTTTATT	TACTTACTTT	8330
	CTGAGACAGA	GTTATGCTCT	TGTTGCCAGG	GCTGGAGTGC	AGCGGCATGA	TCTTGCTCTA	CTGCAACCTC	8400
	CGTCTCCTGG	GTTCAAGCAA	TTCTCGTGCC	TCAGCTCCCC	AAGTAGCTGG	GATTTACAGG	GTGACACCAC	8470
50	ACACCCGGCT	AATTTTGTAT	TTTTAGTAGA	GATGGGCTTT	CACCATGTTG	GTCAAGCTGA	TCTCAAAATC	8540
	CTGACCTCAG	GTGATCCGCC	CACCTCAGCC	TCCCAAAGTG	CTGGGATTAC	AGGCATGAGC	CATGCACTT	8610
	GGCTATTATT	ACCATTTTAA	AACCTCCCTG	GGCTCAAGTC	ACACCCACTG	GTAAGGAGTT	CTGGAGTTTC	8680
	AATTTCCCTT	TACTCAGGA	GTTACCTTCC	TTTGATATTT	TCTGTAATTC	TTCGTAGACT	GGGGATACAC	8750
	CGTCTCTTGA	CATATTCACA	GTTTCTGTGA	CCACCTGTGA	TCCCATGGGA	CCCACCTGCA	GGGCAGCTGG	8820
55	GAGGCTGCAG	GGTTCAGGTC	CCAGTGGGCT	TGCCATCTGC	CAGTAGAAAC	CTGATGTAGA	ATCAGGGCGC	8890
	AAGTGTGGAC	ACTGTCTCTG	ATCTCAATGT	CTCAGTGTGT	GCTGAAACAT	GTAGAAATTA	AAGTCCATCC	8960
	CTCCTACTCT	ACTGGGATTG	AGCCCCTTCC	CTATCCCCCC	CCAGGGGCAG	AGGAGTTCCT	CTCACTCCTG	9030
	TGGAGGAAGG	AATGATACTT	TGTTATTTTT	CAGCTGTGGT	ACTGAATCCA	CTGTTTCATT	TGTTGGTTTG	9100
	TTTGTTTTGT	TTTGAGAGGC	GGTTTCACTC	TTGTTGCTCA	GGCTGGAGGG	AGTGCATATG	CGCGATCTTG	9170
60	GCTTACTGCA	GCCTCTGCCT	CCCAGGTTCA	AGTGATTCTC	CTGCTTCCGC	CTCCCATTTG	GCTGGGATTA	9240
	CAGGCACCCG	CCACCATGCC	CAGCTAATTT	TTTGATTTTT	TAGTAGAGAC	GGGGGTGGGT	GAGGGTTACC	9310
	ATGTTGGCCA	GGCTGGTCTC	GAACCTCTGA	CCTCAGATGA	TCCACCTGCC	TCTGCCTCCT	AAAGTGCTGG	9380
	GATTACAGGT	GTGAGCCACC	ATGCCACGCT	CAGAAATTTAC	TCTGTTTAGA	AACATCTGGG	TCTGAGGTAG	9450
65	GAAGCTCACC	CCACTCAAGT	TGTGTGGTGT	TTTAAGCCAA	TGATAGAATT	TTTTTATTGT	TGTTAGAAAC	9520
	CTCTTGATGT	TTTACACTGT	GATGACTAAG	ACATCATCAG	CTTTTCAAAG	ACACACTAAC	TGCACCCATA	9590
	ATACTGGGTT	GTCTTCTGGG	TATCAGCAAT	CTTCATTGAA	TGCCGGGAGG	CGTTTCCCTG	CCATGCACAT	9660
	GGTGTTAATT	ACTCCAGCAT	AATCTTCTGC	TTCCATTCTT	TCTCTTCCCT	CTTTTAAAT	TGTTGTTTCT	9730
	ATGTTGGCTT	CTCTGCAGAG	AACCAGTGTA	AGCTACAAC	TAACCTTTGT	TGGAACAAAT	TTTCCAAACC	9800
	GCCCCTTTGC	CCTAGTGGCA	GAGACAATTC	ACAAACACAG	CCCTTTAAAA	AGGCTTAGGG	ATCACTAAGG	9870
70	GGATTCTAG	AAGAGCGACC	TGTAATCCTA	AGTATTTACA	AGACGAGGCT	AACCTCCAGC	GAGCGTGACA	9940
	GCCCAGGGAG	GGTGCAGGCG	CTGTTCAAAT	GCTAGCTCCA	TAAATAAAGC	AATTTCTCTC	GGCAGTTTCT	10010
	GAAAGTAGGA	AAGGTTACAT	TTAAGGTTGC	GTTTGTGTAG	ATTTACAGTG	TTGCCGACCT	CAGCTACAGT	10080
	ATCCCTGCAA	GGCCTCGGGA	GACCCAGAAG	TTTCTCGCCC	CCTTAGATCC	AAACTTGAGC	AACCCGACGC	10150
	CTGGATTCTT	GGGAAGTCCT	CAGCTGTCTT	GCGGTTGTGC	CGGGGCCCCA	GGTCTGGAGG	GGACCACTGG	10220
	CCGTGTGGCT	TCTACTGCTG	GGCTGGAAGT	CGGGCTCTCT	AGCTCTGCAG	TCCGAGGCTT	GAGCCAGGCT	10290
75	GCCTGGACCC	CGAGGCTGCC	CTCCACCCTG	TGCGGGCGGG	ATGTGACCAG	ATGTTGGCCT	CTATGCGCAG	10360
	ACAGAGTGCC	GGGGCCACAG	GTCAAGGCCG	TTGTGGCTGG	TGTGAGGCGC	CCGGTGCCTG	GCCAGCAGGA	10430
	GCGCTTGGCT	CCATTTCCTA	CCCTTTCTCG	ACGGGACCGC	CCCGGTGGGT	GATTAAACAG	TTTGGGGTGG	10500
	TTTGCTCATG	GTGGGACACC	CTCGCCGCCT	GAGAACCTGC	AAAGAGAAAT	GACGGGCTTG	TGTCAGGAG	10570

	CCCAAGTCGC	GGGGAAGTGT	TGCAGGGAGG	CACTCCGGGA	GGTCCCGCGT	GCCCGTCCAG	GGAGCAATGC	10640
	GTCTTCGGGT	TCGTCCCCAG	CCGCGTCTAC	GCGCCTCCGT	CCTCCCTTTC	ACGTCCGGCA	TTCTGTGGTG	10710
	CCGGAGCCCG	ACGCCCCGCG	TCCGGACCTG	GAGGCAGCCC	TGGGTCTCCG	GATCAGGCCA	GCGGCCAAAG	10780
5	GGTCGCCGCA	CGCACCTGTT	CCCAGGGCCT	CCACATCATG	GCCCTTCCCT	CGGGTTACCC	CACAGCCTAG	10850
	GCCGATTGCA	CCTCTCTCCG	CTGGGGCCCT	CGCTGGCGTC	CCTGCACCCT	GGGAGCGCGA	GCGGCGCGCG	10920
	GGCGGGGAAG	CGCGGCCAG	ACCCCGGGT	CCGCCCGGAG	CAGCTGCCGT	GTGCGGGCCA	GGCCGGGCTC	10990
	CCAGTGGATT	CGCGGGACA	GACGCCCAGG	ACCGCGCTCC	CCACGTGGCG	GAGGGACTGG	GGACCCGGGC	11060
	ACCCGTCTTG	CCCTTTCACC	TTCCAGCTCC	GCCTCTCCCG	CGCGGACCCC	GCCCCGTCCC	GACCCCTCCC	11130
10	GGGTCCCCGG	CCAGCCCCC	TCCGGGCCCT	CCCAGCCCTT	CCCTTTCCTT	TCCGCGGGCC	CGCCCTCTCC	11200
	TCGCGGCGCG	AGTTTCAGGC	AGCGCTGCGT	CCTGCTGCGC	ACGTGGGAAG	CCCTGGCCCC	GGCCACCCCC	11270
	GGGATGCCGC	GGCTTCCCG	CTGCCGAGCC	GTGCGTCCCT	TGCTGCGCAG	CCACTACCGC	GAGGTGCTGC	11340
	CGCTGGCAC	GTTCGTGCGG	CGCTTGGGCG	CCCAGGGCTG	GCGGCTGGTG	CAGCGCGGGG	ACCCGGGCGC	11410
	TTTCCGCGCG	CTGGTGGCCC	AGTGCCTGGT	GTGCGTGCCC	TGGGACGCAC	GGCCGCCCCC	CGCCGCCCCC	11480
	TCCTTCCGCC	AGGTGGGCCT	CCCCGGGGTG	GGCGTCCGCG	TGGGGTTGAG	GGCGGCCGGG	GGGAACCCAG	11550
15	GACATGCCGA	GAGCAGCGCA	GGCGACTCAT	GGCGCTTCCC	CCGCAAGTGT	CCTGCCTGAA	GAGCTGGGTG	11620
	GCCCGAGTGC	TGCAGAGGCT	GTGCGAGCGC	GGCGCAAGAA	ACGTGCTGGC	CTTCGGCTTC	GCGCTGCTGG	11690
	ACGGGGCCCG	CGGGGGCCCC	CCCCAGGGCT	TCACCACGAG	CGTGCGCAGC	TACCTGCCCA	ACACGGGTGAC	11760
	CGAGCAGCTG	CGGGGAGGCG	GGGCGTGGGG	GCTGCTGTGG	GCGCGCTGGG	GCGACAGCGT	GCTGGTTTAC	11830
20	CTGCTGGCAC	GCTGCGCGCT	CTTTGTGCTG	GTGGCTCCCA	GCTGCGCCTA	CCAGGTGTGC	GGGCGGCCGC	11900
	TGTACCACTG	CGGCGCTGCC	ACTCAGGGCC	GGCCCCCGCC	ACACGCTAGT	GGACCCCGAA	GGCGTCTGGG	11970
	ATGCGAAGCG	GCTTGAAC	ATAGCGTCAG	GGAGGCCGGG	GTCCCCCTGG	GCCTGCCAGC	ACCCGGGTCG	12040
	AGGAGGCGCG	GGGGCAGTGC	CAGCCGAAGT	CTGCCGTTGC	CCAAGAGGCC	CAGGCGTGGC	GCTGCCCTTG	12110
	AGCCGGAGCG	GACGCCCCGT	GGGCGAGGGT	CCTGGGCCCA	CCCGGGCAGG	ACGCGTGGAC	CGAGTGACCG	12180
25	TGGTTTCTGT	GTGGTGTAC	CTGCCAGACC	CGCCGAAGAA	GCCACCTCTT	TGGAGGGTGC	GCTCTCTGGC	12250
	ACGCGCCACT	CCCACCCATC	CGTGGGCCGC	CAGCACCACG	CAGGCCCCCC	ATCCACATCG	CGGCCACCAC	12320
	GTCCCTGGGA	CACGCTTGT	CCCCCGGTGT	ACGCCGAGAC	CAAGCACTTC	CTCTACTCCT	CAGGCGACAA	12390
	GGAGCAGCTG	CGGCCCTCCT	TCTACTCAG	CTCTCTGAGG	CCCAGCCTGA	CTGGCGCTCG	GAGGCTCGTG	12460
	GAGACCATCT	TTCTGGGTTT	CAGGCCCTGG	ATGCCAGGGA	CTCCCCGAG	GTTGCCCCGC	CTGCCCCAGC	12530
30	GCTACTGGCA	AATGCGGCC	CTGTTTCTGG	AGGTGCTTGG	GAACACGCG	CAGTGCCCTC	ACGGGGTGTG	12600
	CCTCAAGACG	CAGTCCCGCC	TGCGAGCTGC	GGTCAACCCA	CGAGCCGGTG	TCTGTGCCCG	GAGAAAGCCC	12670
	CAGGGCTCTG	TGGCGGCCCC	CGAGGAGGAG	GACACAGACC	CCCGTCGCTT	GGTGCAGCTG	CTCCGCCAGC	12740
	ACAGCAGCCC	CTGGCAGGTG	TACGGCTTCC	TGCGGGCCGT	CCTGCGCCGG	CTGGTGCCCC	CAGGCCTCTG	12810
	GGGTCCGAGC	CACAACGAAC	GCCGCTTCCG	CAGGAACACC	AAGAAGTTCA	TCTCCCTGGG	AAGCATGCCC	12880
35	AAGCTCTCGC	TGCAGGAGCT	GACGTGGAAG	ATGAGCGTGC	GGGACTGCGC	TTGGCTGCGC	AGGAGCCAG	12950
	GTGAGGAGGT	GGTGGCCCTG	GAGGGCCGAG	GCCCCAGAGC	TGAATGCAGT	AGGGGCTCAG	AAAAGGGGGC	13020
	AGGCAGAGCC	TGGTCTCCTC	TGTCTCCATC	GTACGCTGGG	CACACGTGGC	TTTTCGCTCA	GAGAGTCGAG	13090
	TGGACACGGT	GATCTCTGCC	TCTGCTCTCC	CTCTGTCCA	GTTTGCATAA	ACTTACGAGG	TTACCTTCA	13160
	CGTTTTGATG	GACACGCGGT	TTCCAGGCGC	CGAGGCCAGA	GCAGTGAACA	GAGGAGGCTG	GGCGCGGCAG	13230
40	TGGAGCCGGG	TTGCCGGCAA	TGGGGAGAA	TGTTTGAAG	CACAGACGCT	CTGGCGAGGG	TGCCCTCAGG	13300
	TTACCTATAA	TCCTCTTCGC	AATTTCAAGG	GTGGGAATGA	GAGGTGGGGA	CGAGAACCCC	CTCTTCTCGG	13370
	GGGTGGGAGG	TAAGGGTTTT	GCAGGTGCAC	GTGGTCAGCC	AATATGCAGG	TTTGTGTTTA	AGATTTAATT	13440
	GTGTGTTGAC	GGCCAGGTGC	GGTGGCTCAC	GCCGGTAATC	CCAGCACTTT	GGGAAGCTGA	GGCAGGTGGA	13510
	TCACCTGAGG	TCAGGAGTTT	GAGACCAGCC	TGACCAACAT	GGTGAAACCC	TATCTGTACT	AAAAATACAA	13580
45	AAATAGCTG	GGCATGGTGG	TGTGTGCCCT	TAATCCCAGC	TACTTGGGAG	GCTGAGGCGA	GAGAATCACT	13650
	GAAACTCTGT	CTTTAAAAAA	AAAAAGTGT	CGTTGATTGT	GCCAGGACAG	GGTAGAGGGA	GGGAGATAAG	13720
	ACTGTTCTCC	AGCACAGATC	CTGGTCCCCA	CTTTAGGTAT	GAAGAGGGCC	ACATGGGAGC	AGAGGACAGC	13790
	AGATGGGCTC	ACCTGCTGAG	GAAGGGACAG	TGTTTGTGGG	TGTTTGTGGG	ATGGTGCTGC	TGGGCCCTGC	13860
50	CGTGTCCCCA	CCCTGTTTTT	CTGGATTTGA	TGTTGAGGAA	CCTCCGCTCC	AGCCCCCTTT	TGGCTCCGAG	13930
	TGCTCCAGAG	CCCTACCGTG	GCAGCTAGAA	GAAGTCCCGA	TTTCAACCCC	TCCCCACAAA	CTCCCCAAGC	14000
	ATGTAAGACT	TCCGGCCATG	CAGACAAGGA	GGGTGACCTT	CTTGGGGCTC	TTTTTTTCTT	TTTTTTCTTT	14070
	TTATGGTGGC	AAAAGTCATA	TAACATGAGA	TTGGCACTCC	TAACACCGTT	TTCTGTGTAC	AGTGCAGAA	14140
	TGCTAACTCG	GCGGTGTTTA	CAGCAGGTTG	CTTGAAATGC	TGCGTCTTGC	GTGACTGGAA	GTCCCTACCC	14210
55	ATCGAACC	AGCTGCCCTA	CACCTGCTGC	GGCTCAGGTG	GACCAACGCG	AGTCAGATAA	AGCTCATGCA	14280
	ACCCAGTTTT	GCTTTTTGTG	CTCCAGCTTC	CTTCGTTGAG	GAGAGTTTGA	GTCTCTGAT	CAGGACTCTG	14350
	CGTGTCAATG	CTGTTCTCTG	ACTTCAGATG	AGGTCACAAT	CTGCCCCCTG	CTTATGCAGG	GAGTGAGGCG	14420
	TGGTCCCCGG	GTGTCCCTGT	CACGTGCAGG	GTGAGTGAGG	CGTTGCCCCC	AGGTGTCCCT	GTACAGTGTG	14490
	GGGTGAGTGA	GGCGCGGCC	CCGGGTGTCC	CTGTCCCGTG	CAGCGTGATT	GAGGTGTGGC	CCCCGGGTGT	14560
60	CCCTGTCCAG	TGTAGGGTGA	GTGAGGCGCC	ATCCCCGGGT	GTCCCTGTCA	CGTGTAGGGT	GAGTGAGGCG	14630
	TGGTCCCCGG	GTGTCCCTGT	CCCGTGCAGG	GTGAGTGAGG	CACGTGCCCC	GGGTGTCCCT	GTACAGTGTG	14700
	GGGTGAGTGA	GGCGCGGTCC	CCGGGTGTCC	CTCTCAGGTG	TAGGGTGAGT	GAGGCGCGCG	CCCAGGGTGT	14770
	CCCTGTCCAG	TGTAGGGTGA	GTGAGGACCA	GTCCCTGGGT	GTCCCTCCCC	GGTATAGGGT	GAGTGAGGCA	14840
	CTGTCCCCGG	GTGTCCCTGT	CACGTGCAGG	GTGAGTGAGG	GCGGGCCCCC	GGGTGTCCCT	CTCAGGTGCA	14910
65	GGGTGAGTGA	GGCGCTGTCC	CTGGGTGTCC	CTGTCTCGTG	TAGGGTGAGT	GAGGCTCTGT	CCCCAGGTGT	14980
	CCTTGGCGTT	TGCTCACTTG	AGCTTGTCTC	TGAATGTTTG	CTCTTTCTAT	AGCCACAGCT	GCGCCGGTTG	15050
	CCCATTTGCT	GGTAGATGG	TGCAGGCCGA	GTGCTGGTCC	CCAAGCCTAT	CTTTTCTGAT	GTCTGGTCTG	15120
	TCTTGGTCA	CTCTCCGTTT	CATTTTGTCT	CGGGACACAG	GGACTGCAGG	CTCTCGCTC	CCGCGTGCCA	15190
	GGCATGTGAG	CCACAGCTTC	AGGTCCGCTT	GCCTCTGTTG	GGCCTGGCTT	GCTCACCACG	TGCCCCGCCA	15260
70	ATGCATGGTG	CCAATACTCC	TCTCCAGCT	TGCTTCAATG	CGAGGCTGGA	CTCTGGGGTG	CTGTGTCTG	15330
	CTGCCACGTG	TTGCTGGAGA	CATCCAGAAA	AGGGTCTCTT	GTGCCCTGAA	GGAAAGCAAG	TCACCCAGC	15400
	CCCTCACTTT	GTCTGTTTTT	CTCCCAAGCT	GCCCTCTGCG	TTGGCCCCCT	TGGGTGGGTG	GCAACGCTTG	15470
	TCACCTTATT	CTGGGCACCT	GCGGCTCAT	GCTTAGGCTG	GGCTCTGCTT	CCAGTCGCCC	CCTCATCTGG	15540
	ATTGACGTCC	AGCCACAGGT	TGGAGTGTCT	CTGTCTGTCT	CCTGCTCTGA	GACCCACGTG	GAGGGCCGGT	15610
75	GTCTCCGCGA	GCCTTCGTC	GACTTCCCTC	TTGGGTCTTA	GTTTTGAATT	TCACTGATTT	ACCTCTGACG	15680
	TTTCTATCTC	TCCATTGTAT	GCTTTTTCTT	GGTTTATCTT	TTTATCTTCT	TTTATGCTTC	TTAGTTTATG	15750
	CATGCCCTTC	CCTCTAAGTG	CTGCCCTTACC	TGCACCCCTG	GTTTTGATGT	GAAGTAATCT	CAACATCAGC	15820
	CACCTTCAAG	TGTTCTTAAA	ATACTTCAAA	GTGTAAATAC	TTCTTTTAA	TATTTCTTAT	CTGTGATTTT	15890
	TTTCTTTGTG	CACGCTGTGT	TTTGACGTGA	AATCAATTTG	ATATCAGTGA	CTTTTAAAGTA	TTCTTTAGCT	15960

	TATTCTGTGA	TTTCTTTGAG	CAGTGAGTTA	TTTGAACACT	GTTTATGTTT	AAGATATGTA	GAGTATCAAG	16100
	ATACGTAGAG	TATTTTAAAGT	TATCATTTTA	TTATTGATTT	CTAACTCAGT	TGTGTAGTGG	TCTGTATAAT	16170
	ACCAATTATT	TGAAGTTTGC	GGAGCCTTGC	TTTGTGATCT	AGTGTGTGCA	TGGTTTCCAG	AACTGTCCAT	16240
5	TGTAAATTTG	ACATCCTGTC	AATAGTGGGC	ATGCATGTTT	ACTATATCCA	GCTTATTAAG	GTCCAGTGCA	16310
	AAGCTTCTGT	CTCCTTCTAG	ATGCATGAAA	TTCCAAGAAG	GAGGCCATAG	TCCCTCACCT	GGGGGATGGG	16380
	TCTGTTTCA	TCTTCTCGTT	TGGTAGCATT	TATGTGAGGC	ATTGTTAGGT	GCATGCACGT	GGTGAATTT	16450
	TTATCTTCCT	GATGAGTGAA	TCTTTTGGAG	ACTTCTATGT	CTCTAGTAAT	CTAGTAATTC	TTTTTTTAAA	16520
	TTGCTCTTAG	TACTGCCACA	CTGGGCTTCT	TTTGATTAGT	ATTTTCTGCT	TGTGTCTGTT	TTCTGCCTTT	16590
10	AATTTATATA	TATATATATA	TTTTTTTTTT	TTTTGAGACA	GAGTCTTGGT	CTGTGCGCCA	GGGTGAGTGC	16660
	AGTGGTGTGA	TCACAGGTCA	GTGTAACCTT	TACCTTCTGG	CCTGAGCCGT	CCTCTCACCT	CAGCCTCCTG	16730
	AGTAGCTGGA	ACTGCAGACA	CGCACCGCTA	CACCTGGCTA	ATTTTAAAT	TTTTTCTGGA	GACAGGGTCT	16800
	TGGCTGTGTG	CCCAGGCTGG	TCTCAAACCT	TTGGACTCAA	GGGATCCATC	TACCTCGGCT	TCCCAAAGTG	16870
	CTGAATTACA	GGCATGAGCC	ACCATGTCTG	GCCTAATTTT	CAACACTTTT	ATATTCTTAT	AGTGTGGGTA	16940
	TGTCCTGTGA	ACAGCATGTA	GGTGAATTTT	CAATCCAGTC	TGACAGTCGT	TGTTAACTG	GATAACCTGA	17010
15	TTTATTTTCA	TTTTTTTGTG	ACTAGAGACC	CGCCTGGTGC	ACTCTGATT	TCCACTTGCC	TGTTGATCCT	17080
	CCTCGTTCCT	TTGTTTCTCA	CCACCTCTTG	GGTTGCCATG	TGCGTTTCTT	GCCGAGTGTG	TGTTGATCCT	17150
	CTCCTTGCTT	CCTGGTCACT	GGGCAATTTG	TTTTATTTCT	CTTTGCTTAG	TGTTAGCCCT	TGATCTTTTT	17220
	ATTGTCGTTG	TTTGCTTTTG	TTTATTGAGA	CAGTCTCACT	CTGTACCCCA	GGCTGAGGTC	TAAATGGACA	17290
20	ATCTCGGCTC	ACTGCAACCT	CTGCCCTCCT	GGTTCAAGCA	GTTTCTCATT	CTCAACCTCA	TGAGTAGCTG	17360
	GGATTACAGG	CGCCACCAC	CACGCTGGC	TAATTTTGTG	ATTTTTAGTA	GAGATAGGCT	TTCACCATGT	17430
	TGGCCAGGCT	GGTCTCAAAC	TCTGACCTG	AAGTAGCTCA	CCCGCTTGG	CCTCCACAG	CCTGCGGATT	17500
	ACAGGTGCAA	GCCACCGTGC	CCGGCATACC	TTGATCTTTT	AAAATGAAGT	CTGAAACATT	GCTACCCCTG	17570
	TCCTGAGCAA	TAAGACCCTT	AGTGATTTTT	AGCTCTGGCC	ACCCCCACAG	CTGTGTGCTG	TTTTCCCTGC	17640
25	TGACTTAGTT	CTATCTCAGG	CATCTTGACA	CCCCACAAG	CTAAGCATT	TTAATATTGT	TTTCCGTGTT	17710
	GAGTGTCTT	GTAGCTTTGC	CCCCGCCCTG	CTTTTCTCTC	TTTGTTCCTC	GTCTGTCTTC	TGCTCTCAGG	17780
	CCGCGCTGCT	GGGTCCCTTT	CCTTGTCTCT	TGCGTGGTTC	TTCTGTCTTG	TTATTGCTGG	TAAACCCACG	17850
	CCTTACCTGT	GCTGGCTCTC	ATGGCATCTA	GCGACCTCCG	GGGACCTCTG	CCTATGATGC	ACAGATGAAG	17920
	ATGTGGAGAC	TCACGAGGAG	GGCGGTCTAT	TTGGCCCGTG	AGTGTCTGGA	GCACCCAGTG	GCCAGCGTTC	17990
30	CCTAGCCAGT	GAGTGACAGC	AACGTCCCGT	CGGCTGGGTT	TCAGCCTGGA	AAACCCACAG	CATGTGCGGG	18060
	TCTGGTGGCT	CCGCGGTGTC	GAGTTTGAAA	TCGCGCAAA	CTGCGGTGTC	GCGCCAGCTC	TGACGGTGCT	18130
	GCCTGGCGGG	GGAGTGTCTG	CTTCTCTCCT	TCTGCTTGGG	AACCAGGACA	AAGGATGAGG	CTCCGAGCCG	18200
	TTGTGCGCCA	ACAGGAGCAT	GACGTGAGCC	ATGTGGATAA	TTTTAAATTT	TCTAGGCTGG	GCGCGGTGGC	18270
	TCACGCTGCT	AATCCCAGCA	CTTTGGGAGG	CCAAGGCGGG	TGGATCACGA	GGTCAGGAGG	TCGAGACCAT	18340
35	CCTGGCCAA	ATGATGAAAC	CCCCTCTGTA	CTAAACACAC	AAAAATTAGC	TGGCGCTGGT	GCGCGGTGCC	18410
	TGTAATCCCA	GCTACTCGGG	AGGCTGAGGC	AGGAGAAATG	CCTGAACCTG	GGAGTTGGAA	GTTGCAAGTA	18480
	GCCGACATTG	CACCACTGCA	CTCCAGCTCG	GCAACACAGC	GAGACTCTGT	CTCAAAAAAA	AAAAAAAAAA	18550
	AAAAAATAAA	AATTTAGTA	GCCACATTAA	AAAAATAAAA	AAGAAAAAGT	GAAATTAATG	TAATAATAGA	18620
	TTTTACTGAA	GCCACAGCAT	TCCACACCTC	ATCATTTTAT	GGTGTATTG	GTGGGAGCAT	CACCTCACAG	18690
40	ACATTTGACA	TTTTTTGAGC	TTTGTCTGCG	GGATCCCCTG	TGTAGGTCCC	GTGCGTGGCC	ATCTCGGCCT	18760
	GGACCTGCTG	GGCTTCCCAT	GGCCTAGGCT	GTTGTACCAG	ATGGTGCAGG	TCCGGGATGA	GGTCGCCAGG	18830
	CCCTCAGTGA	GCTGGATGTC	CAGTGTCCGG	ATGGTGCACG	TCTGGGATGA	GGTCGCCAGG	CCCTGCTGTC	18900
	AGCTGAGTGT	GTGGTGTCTG	GATGGTGCAG	GTCAGGGGTC	AGGTCTCCAG	GCCCTCGGTC	AGCTGGAGTC	18970
	ATGGAGTCCG	GATGATGCAG	GTCCGGGGTG	AGGTGCGCCG	GCCCTGCTGT	GAGCTGGATG	TGTGGTGTCT	19040
45	GGATGGTGCA	GGTCAGGGGT	GAGGTCTCCA	GGCCCTCGGT	AAGCTGGAGG	TATGGAGTCC	GGATGATGCA	19110
	GGTCCGGGGT	GAGGTCCGCA	GGCCCTGCTG	TGAGTGGTGT	GTGTGGTGTG	TGGATGGTGT	AGGTGTGGGG	19180
	TGAGGTACCC	AGGCCCTGCG	GTGAGCTGGG	TGTGCGGTGT	CTGGATGGTG	CAGGTCTGGA	GTGAGGTCCG	19250
	CAGACGGTGC	CAGACCATGC	GGTGAGCTGG	ATATGCGGGT	TCCGGATGTT	GCAGGTCTGG	GGTGAGGTGT	19320
	CCAGGCCCTG	CTGTGAGTTG	GATGTGGGGT	GTCCGGATGC	TGCAGGTCCG	GTGTGAGGTC	ACCAGGCCCT	19390
50	GCTGTGAGCT	GGATGTGTGG	TGTCTGGATG	GTGCAGGTCT	GGGGTGAAGG	TCGCCAGGCC	CCTGCTTGTG	19460
	AGCTGGATGT	GTGGTGTCTG	GATGGTGCAG	GTCTGGAGAT	AGGTGCGCCG	GCCCTCGGTC	AGCTGGATGT	19530
	GCAGTGTCCA	GATGGTGCAG	GTCCGGGGTG	AGGTGCGCCG	ACCTGCGGTC	GAGCTGGATG	TGCGGTGTCT	19600
	GGATGGTGCA	GGTCTGGAGT	GAGGTGCGCA	GGCCCTCGGT	GAGCTGGATG	TATGGAGTCC	GGATGGTGCC	19670
	GGTCCGGGGT	GAGGTGCGCA	GACCTCTGCT	TGAGCTGGAT	GTGCGGTGTC	TGGATGGTAC	AGGTCTGGAG	19740
55	TGAGGTGCGC	AGACCTGTCT	GTGAGCTGGA	TATGCGGTGT	CCGATGGTGT	CAGGTGAGGT	GTGAGGTCTC	19810
	CAGGCCCTCG	GTGAGCTGGA	GGTATGGAGT	CCGATGATG	CAGGTCCGGG	GTGAGGTGCG	CAGGCCCTGC	19880
	TGTGAACCTG	ATGTGCGGGC	TCTGGATGTT	GCAGGTCTGG	GGTGTGGTGC	CCAGGCCCTC	GGTGAGCTGG	19950
	AGGTATGGAG	TCCGGATGAT	GCAGGTCCGG	GGTGAGGTCT	CCAGGCCCTG	CTGTGAGCTG	GATGTGCGGC	20020
	GTCTGGATGG	TGCAGGTCTG	GGGTGTGGTC	GCCAGGCCCT	CGGTGAGCTG	GAGGTATGGA	GTCCGGATGA	20090
60	TGCAGGTCCG	GGGTGAGGTT	GCCAGGCCCT	GCTGTGAGCT	GGATGTGCTG	TATCCGGATG	GTGCAAGTCC	20160
	GGGTGAGGTC	GCCAGGCCCT	GCTGTGAGCT	GGATGTGCTG	TATCCGGATG	GTGCAAGTCT	GGGTGAGGTC	20230
	CACCAGGCC	TGCGGTGAGC	TGGTTGTGCG	GTGTCCGGTT	GCTGCAGGTC	CGGGGTGAGT	TCGCCAGGCC	20300
	CTCGGTGAGC	TGGATGTGCG	GTGTCCCGCT	GTCCGGATGG	TGCAGGTCCA	GGGTGAGGTC	GCTAGGCCCT	20370
65	TGGTGGGCTG	GATGTGCCGT	GTCCGGATGG	TGCAGGTCTG	GGGTGAGGTC	GCCAGGCCCT	TGTTGAGCTG	20440
	GATGTGCCGT	GTCTGCATGG	TGCAGGTCTG	GGGTGAGGTC	GCCAGGCCCT	TGGTGGGCTG	GATGTGTGGT	20510
	GTCCGGATGG	TGCAGGTCCG	GCGTGAGGTC	GCCAGGCCCT	GCTGTGAGCT	GGATGTGCGG	TGCTTGGATG	20580
	GTGCAAGTCC	GGGTGAGGTC	AGCCAAGGCC	TTCGGTGAGC	TGGATGTGGG	GTGTCCGGAT	GTGCAAGGTC	20650
	CGGGGTGAGG	TCGCCAGGCC	CTGCGGTTAG	CTGGATATGC	GGGTGCCGGA	TGGTGCAGGT	CCGGGGTGAG	20720
	GTCAACAGGC	CCTGCGGTTA	GCTGGATGTC	CGGTGTCTGG	ATGGTGCAGG	TCCGGGGTGA	GGTCGCCAGG	20790
70	CCCTGCTGTG	AGCTGGATGT	GCTGTATCCG	GATGGTGCAG	GTCCGGGGTG	AGGTGCGCAG	GCCCTGCAGT	20860
	GAGCTGGATG	TGCTGTATCC	GGATGGTGCA	GGTCTGGCGT	GAGGTGCGCA	GGCCCTGCGG	TTAGCTGGAT	20930
	ATCGGGTGTG	GGATGGTGCA	GGTCCGGGGT	GAGGTACCCA	GGCCCTGCGG	TTAGCTGGAT	GTGCGGTGTC	21000
	CGGATGGTGT	AGGTCTGGGG	TGAGGTCCGC	AGGCCCTGCT	GTGAGCTGGA	TGTGCTGTAT	CCGGATGGTG	21070
75	CAGGTCCGGG	GTGAGGTCCG	CAGGCCCTGC	GGTGAGCTGG	ATGTGCTGTA	TCCGGATGGT	GCAGGTCTGG	21140
	CGTGAGGTGT	CCAGGCCCTG	CGGTGAGCTG	GATGTGCAGT	GTACGGATGG	TGCAGGTCCG	GGGTGAGGTC	21210
	GCCAGGCCCT	CGGTGGGGCT	GTATGTGTGT	TGTCTGGATG	GTGCAAGTCC	GGGGTGGAGT	GCCAGGCCCT	21280
	TGCGGTGAGC	TGGATGTGTG	GTGTCTGGAT	GCTGCAGGTC	CGGGGTGAGT	TCGCCAGGCC	CTCGGTGAGC	21350
	TGGATATGCG	GTGTCCCGCT	GTCCGAATGG	TGCAGGTCCA	GGGTGAGGTC	GCCAGGCCCT	TGGTGGGCTG	21420
	GATGTGCCGT	GTCCGGATGG	TGCAGGTCTG	GGGTGAGGTC	GCCAGGCCCT	TGGTGGGCTG	GATGTGCGGT	21490

	GTCCGGATGG	TGCAGGTCCG	GGGTGAGGTC	ACCAGGCCCT	CGGTGATCTG	GATGTGGCAT	GTCTTCTCG	21560
	TTTAAGGGGT	TGGCTGTGTT	CCGGCCGAG	AGCACCGTCT	CGGTGAGGAG	ATCCTGGCCA	AGTTCTTGCA	21630
	CTGGCTGATG	AGTGTGTACG	TCGTCGAGCT	GCTCAGGTCT	TTCTTTTATG	TCACGGAGAC	CACGTTTCAA	21700
5	AAGAACAGGC	TCTTTTCTA	CCGGAAGAGT	GTCTGGAGCA	AGTTGCAAAG	CATTGGAATC	AGTACTGTGA	21770
	TCCCCACGCC	AGGCCTCTGC	TTCTCGAAGT	CCTGGAACAC	CAGCCCGGCC	TCAGCATGCG	CCTGTCTCCA	21840
	CTTGCCTGTG	CTTCCCTGGC	TGTGCAAGTC	TGGGCTGGGA	GCCAGGGGCC	CCGTACACAG	CCTGGTCCAA	21910
	GTGGATTCTG	TGCAAGGCTC	TGACTGCCTG	GAGCTCACGT	TCTCTTACTT	GTAAATCAG	GAGTTTGTGC	21980
	CAAGTGGTCT	CTAGGGTTTG	TAAAGCAGAA	GGGATTTAAA	TTAGATGGAA	ACACTACCAC	TAGCCTCCTT	22050
10	GCCTTTCCCT	GGGATGTGGG	TCTGATTCTC	TCTCTCTTTT	TTTTTTCTTT	TTTGAGATGG	AGTCTCACTC	22120
	TGTTGCCAG	GCTGGAGTGC	AGTGGCATAA	TCTTGGCTCA	CTGCAACCTC	CACCTCCTGG	GTTTAAGCGA	22190
	TTCAACAGCC	TCAGCCTCCT	AAGTAGCTGG	GATTACAGGC	ACCTGCCACC	ACGCCTGGCT	AATTTTTTGA	22260
	CTTTTAGGAG	AGACGGGGTT	TCACCATGTT	GGCGAGGCTG	GCTCTGAACT	CATGACCTCA	GGTGATCCAC	22330
	CCACCTTGGC	CTCCCAAAGT	GCTGGGTTTA	CAGGCTAAGC	CACCGTGCCC	AGCCCCGAT	TCTCTTTTAA	22400
15	TTATGCTGTT	TCTGTATGAA	TCTTCAATCT	ATTGGATTTA	GGTCATGAGA	GGATAAAATC	CCACCCACTT	22470
	GGCGCATGAC	TGCAGGGAGC	ACCTGTGCAG	GGAGCACTG	GGGATAGGAG	AGTTCCACCA	TACGCTAACT	22540
	TCTAGGTGGC	TGCATTTGAA	TGGCTGTGAG	ATTTTGTCTG	CAATGTTCCG	CTGATGAGAG	TGTGAGATTG	22610
	TGACAGATTG	AAGCTGGATT	TGCATCAGTG	AGGGACGGGA	GCCTGGTCT	GGGAGATGCC	AGCCTGGCTG	22680
	AGCCAGGCC	ATGGTATTAG	CTTCTCCCTG	TCCCGCCAG	GCTGACTGTG	GAGGGCTTTA	TACGCTAACT	22750
20	CAGGGCTTCC	CCAGCTCCCC	TGCACACTCG	AGTCCCTGGG	GGGCTTGTG	ACACCCCATG	CCCCAAATCA	22820
	GGATGCTGTC	AGAGGGAGCT	GGCAGCAGAC	CTCGTCAGAG	GTAAACACAG	CTCTGGGCTG	GGGACCCCGA	22890
	CGTGGTGTG	GGGCCATTTC	CTTGCATCTG	GGGAGGGCTC	AGGGCTTTCC	CTGTGGGAAC	GGTGAATAC	22960
	ACAATGCACC	TTACTTAGAC	TTTACACGTA	TTTAATGGTG	TGCGACCCAA	CATGGTCATT	TGACCAAGTAT	23030
	TTTGGAAAGA	ATTTAATTGG	GGTGACCCGA	AGGAGCAGAC	AGACGTGGTG	GTCCCCAAGA	TGCTCCTTGT	23100
25	CACTACTGGG	ACTGTTGTTT	TGCTTGGGGG	GCCTTGGAGG	CCCTCCTCTC	CTGGACAGGG	TACGCTAACT	23170
	TTTCTACTCT	GCTGGGCCCT	CGGCCTGCGG	TCAGGGCACC	AGCTCCGGAG	CACCCGCGGC	CCCAGTGTCC	23240
	ACGGAGTGCC	AGGCTGTGAG	CCACAGATGC	CCAGGTCAG	GTGTGGCCGC	TCCAGCCCCC	TGCCCCCAT	23310
	GGGTGTTTTT	GGGGGAAAAG	GCCAAGGGCA	GAGGTGTGAG	GAGACTGGTG	GGCTCATGAG	AGCTGATTCT	23380
	GCTCCTTGGC	TGAGCTGCCC	TGAGCAGCCT	CTCCCGCCCT	CTCCATCTGA	AGGGATGTGG	CTCTTTCTAC	23450
30	CTGGGGGTCC	TGCTTGGGGG	CAGCCTTGGG	CTACCCCATG	GGCTGTACCA	GAGGGACAGG	CATCCTGTGT	23520
	GGAGGGCAT	GGGTTACAGT	GGCCCCAGT	GCAGGCTGGG	ACCAGGCTCC	CTGGTGCTGA	TGGTGGGACA	23590
	GTCAACCTGG	GGGTTGACCG	CCGGACTGGG	CGTCCCCAGG	GTGACTATA	GGACCAAGTG	TCCAGGTGCC	23660
	CTGCAAGTAG	AGGGGCTCTC	AGAGGCGTCT	GGCTGGCATG	GGTGGACGTG	GCCCCGGGCA	TGGCCTTCAG	23730
	CGTGTGCTGC	CGTGGGTGCC	CTGAGCCCTC	ACTGAGTCTG	TGGGGGCTTG	TGGCTTCCCG	TGAGCTTCCC	23800
35	CCTAGTCTGT	TGTCTGGCTG	AGCAAGCCTC	CTGAGGGGCT	CTCTATTGCA	GACAGCACTT	GAAGAGGGTG	23870
	CAGCTGCGGG	AGCTGTGCGA	AGCAGAGGTC	AGGCAGCATC	GGGAAGCCAG	GCCCCGCTCT	CTGACGTCCA	23940
	GACTCCGCTT	CATCCCCAAG	CTGACGGGCT	TGCGGCCGAT	TGTGAACATG	GACTACGTCT	TGGGAGCCAG	24010
	AACGTTCCGC	AGAGAAAAGA	GGGTGGCTGT	GCTTTGGTTT	AACTTCCTTT	TTAAACAGAA	GTGCGTTTGA	24080
	GCCCCACATT	TGGTATCAGC	TTAGATGAAG	GGCCCCGAGG	AGGGGCCACG	GGACACAGCC	AGGGCCATGG	24150
40	CACGGGCGCA	ACCCATTGTT	GGTCACAGTG	AGGTGGCCGA	GGTGCCGGTG	CCTCCAGAAA	AGCAGCGTGG	24220
	GGGTGTAGGG	GGAGCTCCTG	GGGCAGGGAC	AGGCTCTGAG	GACCACAAGA	AGCAGCCGGG	CCAGGGCCTG	24290
	GATGCAGCAC	GGCCCCGAGT	CCTGGATCCG	TGTCTCTGCT	TGGTGCAGAG	CCTCCGTGCG	CTTCCGCTTA	24360
	CGGGGCCCGG	GGACCAAGGC	ACGAGCTCCA	GGGCTCTGAG	GGGCTCTGAG	GATCCTGGAC	CTTGGCCCA	24430
	GGCTCCTGCA	CCCCACCCCT	GTGGCTGCGG	TGGCTGCGGT	GACCCCGTCA	TCTGAGGAGA	GTGTGGGGTG	24500
45	AGGTGGACAG	AGGTGTGGCA	TGAGGATCCC	GTGTGCAACA	CACATGCGGC	CAGGAACCCG	TTTCAAACAG	24570
	GGTGTGAGGA	AGCTGGGAGG	GGTCTAGTGT	CCCGGGTCTG	GGTGGCTGGG	GACACTGGGG	AGGGGCTGCT	24640
	TCTCCCTTGG	GTCCCTATGG	TGGGGTGGGC	ACTTGGCCGG	ATCCACTTTC	CTGACTGTCT	CCCATGCTGT	24710
	CCCCGCCAGG	CCGAGCGTCT	CACCTCGAGG	GTGAAGGCAC	TGTTACAGCT	GCTCAACTAC	GAGCGGGGCG	24780
	GGCGCCCGGG	CCTCCTGGGC	GCCTCTGTGC	TGGGCTGGGA	CGATATCCAC	AGGGCCTGGC	GCACCTTCCT	24850
50	GCTGCGTGTG	CGGGCCCCAG	ACCCGCCGCC	TGAGCTGTAC	TTTGTCAAGG	TGGGTGCCGG	GGACCCCCGT	24920
	GAGCAGCCCT	GCTGGACCTT	GGGAGTGGCT	GCCTGATTGG	CACCTCATGT	TGGGTGGAGG	AGGTACTCCT	24990
	GGGTGGGGCC	CAGGGAGTGC	AGGTGACCTT	GTCACTGTTT	AGGACACACC	TGGCACCTAG	GGTGGAGGCC	25060
	TTACAGCCTT	CCTGCAGCAC	ATGGGGCCGA	CTGTGCACCC	TGACTGCCCG	GGCTCCTATT	CCCAAGGAGG	25130
	GTCCCACTGG	ATTCCAGTTT	CCGTACAGAG	AGGAACCCGA	ACGGCTCAGC	CACCAAGGCC	CGGTGCCTTG	25200
55	CACCCAGCTC	CTGAGCCAGG	GGTCTCCTGT	CCTGAGGCTC	AGAGAGGGGA	CACAGCCCGC	CCTGCCCTTG	25270
	GGGTCTGGAG	TGGTGGGGGT	CAGAGAGAGA	GTGGGGGACA	CCGCCAGGCC	AGGCCCTGAG	GGCAGAGGTG	25340
	ATGTCTGAGT	TTCTGCGTGG	CCACTGTGAG	TCTCCTCGCC	TCCACTCACA	CAGGTGGATG	TGACGGGGCG	25410
	GTACGACACC	ATCCCCCAGG	ACAGGCTCAC	GGAGGCTCAT	GCCAGCATCA	TCAAACCCCA	GAACACGTAC	25480
	TGCGTGCGTC	GGTATGCCGT	GGTCCAGAAG	GCCGCCCATG	GGCACGTCCG	CAAGGCCTTC	AAGAGCCACG	25550
60	TAAGGTTTAC	GTGTGATAGT	CGTGTCCAGG	ATGTGTGTCT	CTGGGATATG	AATGTGTCTA	GAATGCAGTC	25620
	GTGTCTGTGA	TGCGTTTCTG	TGGTGGAGGT	ACTTCCATGA	TTTACACATC	TGTGATATGC	TGTGTGGGCA	25690
	CGTGTGTGTC	GTGGTGCATG	TATCTGTGGC	GTGCATATTT	GTGGTGTGTG	TGTGTGTGGC	ACGTGTGTGT	25760
	CCATGGTGTG	TGTGCTGTGT	GTGTGCTATG	GTGTGTGTCT	GTGACACGTG	CATGTTTCATG	CTGTGTGCTG	25830
	CATGTCTGTG	ATGTGCTTAT	TTGTGGTGTG	TGTGCTGTATG	TGTCCGTGAC	ATATGCTGTG	CTATGGCATG	25900
65	GGTGTGTGTG	GCCCCCTTGG	CTTACTCCTT	CCTCCTCCAG	GCATGGTCCG	CACCAATTGT	CTCAGCTCTT	25970
	CGGGTGTCTG	TTTGGGGAGC	TCCACATTCA	GGGTCTCTAC	TTCTAGCATG	GGTGGCCCTG	TTCTGTCTCA	26040
	GGGCTGGGCG	TTGGAGACTG	TAAGCCAGGT	TTGAGGAGGAG	AGTAGGGATG	CTGGTGGTAC	TGCCCTGGAC	26110
	CCCTGGCACC	CCAGGACCCC	CAGTCTGGCT	TATGCCGGCT	CCATGAGATA	TAGGAAGGCT	GATTACAGGC	26180
	TGCGTCCCCG	GGACACACTC	CTCCAGAGGC	GGCGGGGGCT	CTTGGGGGCT	GGCAGGGGGT	AAAGGGGGCC	26250
	TGGGCTTGGG	TTCCACCCCA	GTGCTCATGA	GCACGCTGGA	GGGGTAAGCC	CTCAAAGTCG	CTCAGGGCCG	26320
70	GGGTGCAGAG	GTGAAGAAGT	ATCCCTGGAG	CTTCCGTCTG	GGGAGAGGCA	CATGTGGAAA	CCCACAAGGA	26390
	CCTCTTTCTC	TGACTTCTTG	AGCT					26414

## Contig 2:

	TGTGGGATTG	GTTTTCATGT	GTGGGATAGG	TGGGGATCTG	TGGGATTGGT	TTTTATGAGT	GGGGTAACAC	70
	AGAGTTCAAG	GCGAGCTTTC	TTCTGTAGT	GGGTCTGCAG	GTGCTCCAAC	AGCTTTATTG	AGGAGACCAT	140
5	ATCTTCCTTT	GAACATATGGT	CGGGTTTATA	GTAAAGTCAGG	GGTGTGGAGG	CCTCCCCTGG	GCTCCCTGTT	210
	CTGTTTCTTC	CACTCTGGGG	TCGTGTGGTG	CCTGCTGTGG	TGTGTGGCCG	GTGGGCAGGG	CTTCCAGGCC	280
	TCCTTGTGTT	CATTGGCCTG	GATGTGGCCC	TGGCTACGCT	CCGTCCCTGG	AATCCCCCTG	CGAGTTGGAG	350
	GCTTCTTTTC	TTTCTTTTTT	TCTTCTTTTT	TTTTTTTTTT	TGATAACAGA	GTCTCGCTCT	TTTTTGCCCA	420
	GGCTGGAGTG	GTTTGGCGTG	ATCTTGGCTC	ACTGCAACCT	GTGCTTCCTG	AGTTCAAGCA	ATTCTCTTGC	490
10	CTCAGCCTCC	CAAGTAGCTG	GAATTATAGG	CGCCACACAC	CATGCTGACT	AATTTTTGTA	ATTTTAGTAG	560
	AGACGAGGTT	TCTCCATGTT	GGCCAGGCTG	GTCTCGAACT	CCTGACCTCA	GGTGATCCCT	CCACCTCGGC	630
	CTCCCAAAGT	GCTGGGATGA	CAGGTGTGAA	CCGCCGCGCC	CGGCCGAGAC	TCGCTTCTCT	CAGCTTCCGT	700
	GAGATCTGCA	GCGATAGCTG	ACCTGCAGCT	TGGTGCTGAC	AACCTCCGTT	TTCTTCTCTC	AGGTCTCGGT	770
	AGGGGTCTTT	CCATTTTCATG	ACTCTCTTCA	CAGAAGAGTT	TCAGTGTGTC	TGATTTCCCG	GCTGTTTCTT	840
	GCGTAATTGG	TGTCTGCTGT	TTATCGATGG	CCTCCTTCCA	TTTCTTTTAG	GCTTTGTTTA	TTGTTGTTTT	910
15	TCCGGTCCCT	TGAAGGAAAA	GTTTCGATTA	TGGATGTTTG	AACCTTCTTT	TCTAAACAAAG	CATCTGAAGT	980
	TGCCGTTTTT	CCTCTAAAGC	AGGGATCCCG	AGGCCCTGG	CTGTGGAGTG	GCACCGGTCT	GGGGCTGTT	1050
	AGGAACCCGG	GCGACAGCGG	GAGGCTAGGT	GGGGTGTGGG	GAGCCAGCGT	TCCCGCCTGA	GCCCCGCCCC	1120
	TCTCAGATCA	GCAGTGGCAT	GCGGTGCTCA	GAGCCGACCA	CACCTACTG	AGAAGTGTGC	GTGAGAGGGG	1190
20	TCTAGATTCT	GTGCTCCTTA	TGGGAATCTA	ATGCCTGATG	ATCTGAGGTG	GAACCGTTTG	CTCCCAAAAC	1260
	CATCCCCCTT	CCCACTGCTG	TCCTGTGGAA	AAATCGTCTT	CCACGAAACC	AGTCCCTGGT	ACCACAATGG	1330
	TTGGGGAAACC	TGTGCTAAAG	ACCTGCTTCA	GCAGCCTCTC	GTCAAGTGTG	ATATATTGGC	TTTTCTGTGT	1400
	TGAGTCCAGA	ATAATTACGG	ATTTCTGTGA	TGCTTTCCGC	CGACCTCAGA	CCCATGGGCT	ATTTGTGGGC	1470
	GTGTTGCTGT	CTCCTGGGTT	GGGAAGGGTG	CAGGCCCATC	GTACCTTCTC	GTTACTGCCT	TCCAGGTTGG	1540
25	TTCTCAGGTT	TGAATCGTAC	TCGATGTGGT	TTTAGCCAC	GGCCCTGCCG	CCAGCTCTCT	GGGGCTGAAG	1610
	AACATGCTGA	AGCACAGAGT	CACCGTGCCT	GTCTTTTGAT	GCCTCACAAG	CTCGAGGCCT	CCTGTGTCCG	1680
	TGTTAGTGTG	TGTCACGTGC	CTGCTCACAT	CCTGTCTTGG	GGACGACAGG	GCTTAGCAGG	TCCCGTAGTA	1750
	AATGACAAGG	GTCTGGGGG	AGTCTGCAGA	ATAGGAGGTG	GGGGTGCCCG	TCTCTCTCCC	GCGCTCTCAG	1820
	ACTCTTCTCC	TGCCTGTGCT	GTGGCTGCAC	CTGCATCCCT	GCAATCCCTC	CAGCACTGGG	CTGGAGAGGC	1890
	CCGGGAGCTG	GAGTGCCACT	TGTGCCACGT	GACTGTGGAT	GGCAGTCCGT	CACGGGGGTC	TGATGTGTGG	1960
30	TGACTGTGGA	TGGCGGTTGG	TCACAGGGGT	CTGATGTGTG	GTGACTGTGT	ATGGCGGTGC	TGGGGTCTGA	2030
	TGTGGTGA	GTGGATGGCG	GTGCTGGGGT	CTGATGTGTG	GTGACTGTGG	ATGGCGGTGC	TGGGGTCTGA	2100
	TGTGGTGA	GTGGATGGCG	GTGCTGGGGT	CTGATGTGTG	GACTGTGGAT	GGCGGTCTGT	GGGTCTGATG	2170
	TGTTGAGTGT	GGATGGCAGT	CGTGGGGTCT	GATGTGTGGT	GACTGTGGAT	GGCGGTCTGT	GGGTCTGATG	2240
35	TGTTGAGTGT	GGATGGCAGT	CGTGGGGTCT	GATGTGTGGT	GACTGTGGAT	GGCGGTCTGT	GGGTCTGATG	2310
	TGTGGTGA	GTGGATGGCG	GTGCTGGGGT	CTGATGTGTG	GTGACTGTGG	ATGGCGGTGC	TGGGGTCTGA	2380
	TGTGTGGTGA	CTGTGGATGG	TGATCGGTCA	CAGGGGTCTG	ATGTGTGGTG	ACTGTGGATG	GCGGTCTGTT	2450
	GGTCTGATGT	GTGGTGA	TGGATGGTGA	TCGGTACAG	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	2520
40	GTGCTGGGTT	GTGGTGA	TGGATGGTGA	TCGGTACAG	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	2590
	GATGGCGATC	GGTCACAGGG	GTCTGATGTG	TGGTGA	GGATGGCGGT	CGTGGGGTCT	GATGTGTGGT	2660
	GACTGTGGAT	GGCGGTCTGT	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	2730
	GACTGTGGAT	GGCGGTCTGT	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	2800
	GACTGTGGAT	GGCGGTCTGT	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	2870
	GACTGTGGAT	GGCGGTCTGT	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	2940
45	GTGGTGA	TGGATGGCAG	TCGTGGGGT	TGATGTGTGG	TGACTGTGGA	TGGCGGTCTG	GGGTCTGATG	3010
	GTGTGGTGA	TGGATGGCAG	TCGTGGGGT	TGATGTGTGG	TGACTGTGGA	TGGCGGTCTG	GGGTCTGATG	3080
	ATGTGTGGTG	ACTGTGGATG	GCGGTCTGTT	GGTCTGATGT	GGTGA	GATGGCGGTG	GTGGGTCTG	3150
	ATGTGTGGTG	ACTGTGGATG	GCGGTCTGTT	GGTCTGATGT	GGTGA	GATGGCGGTG	GTGGGTCTG	3220
	GGGTCTGATG	TGTGGTGA	TGGATGGCAG	TCGTGGGGT	TGATGTGTGG	TGACTGTGGA	GTGGGTCTG	3290
50	GGGTCTGATG	TGTGGTGA	TGGATGGCAG	TCGTGGGGT	TGATGTGTGG	TGACTGTGGA	GTGGGTCTG	3360
	GTACACAGGG	TCTGATGTGT	GGTGA	GATGGCGGTG	GATGGCGGTG	GATGGCGGTG	GTGGGTCTG	3430
	GCGGTCTGATG	GGTCTGATGT	GTGGTGA	TGGATGGCAG	TCGTGGGGT	TGATGTGTGG	TGACTGTGGA	3500
	TGGCGGTCTG	GGGTCTGAT	GTGGTGA	TGGATGGCAG	TCGTGGGGT	TGATGTGTGG	TGACTGTGGA	3570
	GCAAGTGGAG	TCCAGGTTGT	GTCTGTAGCT	ACTTTGCTGT	CTCGGCCCTC	CGGCCCTCTC	TTCCCAAAAC	3640
55	GAAGCTTCCC	AGGCGCTCTC	TGGGCTTTCAT	CCCGCATCG	GGCTTGCCCG	CAGGTCCACA	CTGCTGATC	3710
	GGAAGAAACA	AGTCCCAAGC	TCTGGCCGGG	GCAGGCCACA	TTTGTGGCTC	ATGCCCTCTC	CTCTGCCGGC	3780
	AGGTCTCTAC	CTTGACAGAC	CTCCAGCCGT	ACATGCGACA	GTTCGTGGCT	CACCTGCAGG	AGACCAGCCC	3850
	GCTGAGGAT	GCGGTCTGTC	TCGAGCAGGT	CTGGGCACTG	CCCTGCAGGG	TTGGGCACGG	ACTCCACGCA	3920
	GTGGGTCTCT	CCCTGGGCAA	TCACTGGGCT	CATGACCGGA	CAGACTGTTG	GCCCTGGGGG	GCAGTGGGGG	3990
60	GAATGAGCTG	TGATGGGGG	ATGATGAGCT	GTGTGCCTTG	GCGAAATCTG	AGCTGGGGCA	TGCCAGGCTG	4060
	CGACAGCTGC	TGCATTACAG	CACCTGTCTA	CGTTTGA	GCGGGCTCTC	CTCCAGTTCC	GCAGTGCCTT	4130
	TGTTTCATGAT	TTGCTAAATG	TCTTCTCTGC	CAGTTTTGAT	CTTGAGGCCA	AAGGAAAGGT	GTCCCTCTCC	4200
	TTTAGGAGGG	CAGGCCATGT	TTGAGCCGTG	TCCTGCCACG	CTGGCCCCCT	AGTGTCTGGT	CTAGAGCCAA	4270
	AGGAAACGTG	TCCCTTTCT	TAGGAGGACG	GGCGGTGTTT	GAGCCACGCC	CCGCTGAGCG	GGCTCTCAG	4340
65	TGCTGGGTCT	GTCCACGTGG	CCCTGTGGCC	CTTTGCAGAT	GTGGTCTGTC	CACGTGGCCC	TGTGGCTCTT	4410
	TGCAGATGCC	TGTTAGCACT	TGCTCGGCTC	TAGGGGACAG	TCGTGTCCAC	CGCATGAGGC	TCAGAGACCT	4480
	CTGGGCGAAT	TTCTTTGGCT	CCAGGGTGG	GGGTGTAGGT	GGCTTGGGCT	GCTGGGACCC	AGACCTCTGT	4550
	CCCGGCAGCT	GGGCAGCAAC	TCCTGGATCA	CATATGCCAT	CCGGGCCACG	GTGGGCTGTG	TGGGTGTGAG	4620
	CCAGCTGGA	CCACAGGTTG	GCCAGAGGA	GACGTTCTGT	GTACACACT	CTGCCAAGC	CCATGTGTGT	4690
70	CTGCAGAGA	TCGGCCCGGC	CAGCCACGGA	TGCCCTTCTA	TTCCAGCCCA	GCCCGCAGT	TCCCAACAA	4760
	CACTGACCCC	AAAAGGGACG	GAGGGTCTTG	GCCACGTGGT	CCTGCCTGTC	TCAGCACCCA	CCGGCTCACT	4830
	CCCATGTGTC	TCCCGTCTGC	TTTCGACAG	CTCCTCCCTG	AATGAGGCCA	GCAGTGGGCT	CTTCGACGTC	4900
	TTCTTACGCT	TCATGTGCCA	CCAGCCCGTG	GCATCAGAGG	GCAAGTGA	CAGGTGGGCA	GGTGCCATTG	4970
	CCCTGCCGGT	GGCTGGGCGG	GCTGGCAGGG	CTTCTGCTCA	CCTCTCTCCT	GCCCTTCTCC	CACTGNCCTT	5040
	CTGCCCGGGG	CCACAGAGT	CTCCTTTTCT	GGCCCCCGCC	CCCTCCGGCT	CCTGGGCTGC	AGGCTCCCCG	5110
75	GGCCCCGGA	ACATGGCTCG	GCTTGGCGCA	CCGGAGAGCG	AGCAGGTGCC	ACACGAGGCC	TGGAATGGC	5180
	AAGCGGGTGT	TGGAGTTGCT	CCTGCGTGGA	GGACGAGGGG	CGGGGGTGT	GTCTGGGTCA	GGTGTGCGCC	5250

001250-012500



	GAGCGTTTGA	GCCTGCAGCT	TGTCAGCTCC	AAGTTACTAC	TGACGCTGGA	CACCCGGCTC	TCACACGCTT	5320
	GTATCTCTCT	CTCCCGATAC	AAAAGGATTT	TATCCGATTC	TCATTCTGT	CCCTGTCTGT	TGACCCCGCT	5390
	GAGGGCGCGG	GCTCTTCTCT	CTGTGACTAG	ATTTCCTATC	TGGAAAGTGC	GGGGTTGACC	GTGTAGTTTG	5460
5	CTCCTCTCGG	GGGGCCTGTG	GTGGCCATGG	GGCAGGCGGC	CTGGGAGAGC	TGCCGTCACA	CAGCCACTGG	5530
	GTGAGCCACA	CTCACGGTGG	TAGAGCCACA	GTGCCTGGTG	CCACATCAGC	TCCTCTGGAT	TTTAAGTAAA	5600
	ACCACACACC	TCCCGGCAGG	CATCTGCCTG	CGACCTGTGG	TGTGCTGGGG	GAGAGTGGTA	GCACGGAGGA	5670
	AATTCGTGCA	CACTCAAGGT	CATCAGCAAG	GTCTCCGCA	GTCAAGGTGGA	ACGTGGAGGC	CTCTCTCTGG	5740
	GATCGTCTCC	AGCGGATAAA	GGACTGTGCA	CAGCTTCGGA	AGCTTTTATT	TAAAAATATA	ACTATTAATT	5810
	ATTGCATTAT	AAGTAATCAC	TAATGGTATC	AGCAATTATA	ATATTTTATTA	AAGTATAATT	AGAAATATTA	5880
10	AGTAGTACAC	ACGTCTCTGA	AAAACACAAA	TTGCACATGG	CAGCAGAGTG	AATTTTGGCC	GAGGGACACG	5950
	TGTGCACATG	TGTGTAAGCG	GCCCCCAGGC	CCACAGAATT	CGCTGACAAA	GTCACTCTCC	CAGAGAAGCC	6020
	ACCACGGGCC	TCCTTCGTGG	TCGTGAATTT	TATTAAGATG	GATCAAGTCA	CGTACCGTCC	ACGTGTGGCA	6090
	GGGCTTTGGG	GAATGTGAGG	TGATGACTGC	GTCTCATGTC	CCTGACAGAC	AGGAGGTGAC	TGTGTCTGTC	6160
	CTGTCCCTAG	GACACGGACA	GGCCCCAAGC	TCTAGTCCCC	ATCGTGGTCC	AGTTTGGCCT	CTGAATAAAA	6230
15	ACGTCTTCAA	AACTGTGTTG	CCCCAAAAGT	AAGAACAGAG	AGAGTTTCCC	ATCCCCATGT	CTCACGGAGG	6300
	CGTATCTGCT	TGCGTTGACT	CGCTGGGCTG	GCCGGACTCC	TAGAGTTGGT	GCGTGTGCTT	CTGTGCAAAA	6370
	AGTGCAGTCC	TCTTGCCCAT	CACGTGTGATA	TCTGCACCAG	CAAGGAAAGC	CTCTTTTCTT	TTCTTTCTTT	6440
	TTTTTTTTTT	GAGACGGAAC	GTCACGTGTTG	TCTGCCTGGG	CTTGAGTGCA	GTGGCGGATG	CTCACGGAGG	6510
20	TGCAACCTCC	GCCTCCCGGG	TTCCAGCATT	TCTCTGCTCT	CAGCTCCCG	AGCAGCTGAG	ATTACAGGCA	6580
	CCCACCCCTC	GCGCCTGGCT	AATTTTTGTA	TTTTTAGTAG	AGAGGGGTTT	TTGCCATGTT	GGCCAGGCTG	6650
	GTCTCGAATC	CCTGACCTCA	GGTGATCCAC	CCACCTCGGC	CTCCCAAAGT	GCTGGGATTA	ACGTGTGGCA	6720
	CCATCACGCC	CAGCCGGAAA	GCCTCTTTTT	AAGGTGACCA	CCTATAGCGC	TTCCCGAAAA	TAACAGGTCT	6790
	TGTTTTTGCA	GTAGGCTGCA	AGCGTCTCTT	AGCAACAGGA	GTGGCGTCTC	GTGGGCTCTG	GGGATGGCTG	6860
25	AGGGTCGCTG	GGCAGCCATG	CCTCTGTGTG	GCATCTATTG	GGTCCACGGG	GTTTCTCTGC	CTCACTGTTG	6930
	TGTCTGAAAA	CGCACCCCTT	GCATCCTTGT	TTGGAGAGTT	TCTGCTTCTC	GTTGGTCATG	CTGAAACTAG	7000
	GGGCAAGGTT	GTATCCGTTG	GCGCGCAGCG	GCTACATGTA	GGGTCTAGAG	TCTTTCACCG	TGGACAAATT	7070
	CCTTGAAAAA	AAAAAAAGTA	GTCCGGTTAA	GCATTCATTC	CGGGTCAAGT	GTCTGGTTCT	GTGAATAAAC	7140
	TCTAAGATTT	AAGAAACCTT	AATGAAAGAA	AACCTTGATG	ATTGAGAGCA	AGGATGTGGT	CACACCTGTG	7210
	GCTGGATCTG	TTTCAGCCGC	CCCAGTGCAT	GGTGAGAGTG	GGGAGCAGGG	ATTGTTTGGT	CAGAGGTCTC	7280
30	ATCTGGTATG	TTTCTGAGGT	GTTTGCCGGC	TGAATGGTAT	ACGTGTCTGT	TGTGTGTATG	AGGTTCTGTG	7350
	TCTGTGTGTG	GCTCGGTTTG	AGTGTACGCA	TGTCCAGCAC	ATGCCCTGCC	CGTCTCTCAC	CTGTGTCTTC	7420
	CCGCCCCAGG	TCCTACGTCC	AGTGCCAGGG	GATCCCCGAG	GGCTCCATCC	TCTCCACGCT	GCTCTGCAGC	7490
	CTGTGCTACG	GCGACATGGA	GAACAAGCTG	TTTGCGCGGA	TTCCGGCGGA	CGGGTGAAGG	CTCTCTTCCC	7560
35	CCAGGGGGGC	TTGGGTGGGG	GTTGATTTGC	TTTTGATGCA	TTCACTGTGA	ATATTCCTGG	TGCTCTGGAG	7630
	ACCATGACAG	CTCTGTCTTG	AGGAACCCAG	CAAGCTTGCA	GCCCTTCTTT	GGTATGAAGC	CGCACGGGAG	7700
	GGGTTGACAG	CCCTGAGGAG	TGCCGGCTCC	ACCGAGGCTC	CCAGCTGCTC	TGTCCAGCGG	CCATGCTCAG	7770
	GCTCAGCAGG	CGGGAGGGCC	GCTGCCCTGC	ATGATGAGCA	TGTGAATTCA	ACACCGAGGA	AGCACACCAG	7840
	CTTCTGTCTC	GTCAACCCAG	TTCCGTTAGG	GTCTTGGGG	AGATGGGGCT	GGTGCAGCCT	GAGGCCCCAC	7910
40	ATCTCCAGCA	AGGCCCTCGA	CAGGTGGGCT	GGACTGGGCG	CCTCTTCAGC	CCATTGCCCA	TCCCTCTTGC	7980
	ATGGGGTCTA	CACCCAAGGA	CGCACACACC	TAAATATCGT	GCCAACCTAA	TGTGGTTCAA	CTCAGCTGGC	8050
	TTTTATTGAC	AGCAGTTACT	TTTTTTTTTT	TAATACTTTA	AGTTCTAGGG	TACATGTGCA	AGCAGTGCAG	8120
	GTTAGTAATG	TATGTATACA	TGTGCCATGT	TGGTGTGCTG	CACCCATTAA	CTCATCATTT	CGAATGGTAT	8190
	TATCTCTCTA	TGCTATCCCT	CCCCACTCCC	CCCATCCCAT	GACAGGCCCT	GGTGTGTGAT	GTTCCCCACC	8260
	CTGTGTCCAA	GTGTTCTCAT	TGTTTCAGTT	CCACCTGTGA	GTGAGAACAT	GTGGTGTGTT	GTTTTCTTTC	8330
45	CTTGCAATAG	TTTGCTCAGA	GTGATGGTTT	CCAGCTTCTG	CCATGTCCCT	ACAAAGGACA	TGAACTCATC	8400
	CTTTTTTATG	ACTGCATAGT	ATTCCGTGGT	GTATATGTGC	CACATTTTCT	TAATCCAGTC	TATCATCGAT	8470
	GGACATTTGG	GTTGGTTGCA	AGTCTTTGCT	ACTGTGAATA	GTGCCGCAAT	AAACATACGT	GTCATGTGT	8540
	TTTTATAGCA	GCATGATTTA	TAACTCTTTG	GGTATATACC	CAGTAATGGG	ATGGCTGGGT	TCCCTAAGAG	8610
50	TTCTAGTTCT	AGATCCTTGA	GGAATCACCA	CACGTCTCTC	CACAATGGTT	GAAC TAGTTT	ACACTCCCAC	8680
	CAACAGTGTA	AAAGTGTCTT	GGTGCTGGAG	AGGATGTGGA	CAGCAGTTAT	TTTTTTATGA	AAATAGTATC	8750
	ACTGAACAGA	CAGACAGTTA	GTGAAGGATG	CCTGCAGGAC	ACACAGCCAT	TTCTCTCGAA	TTCTCTCGAA	8820
	GACTCCGGGT	TTTTCTCTGT	CATCTTTTGA	AACCTTAGCT	CCAATTATAG	CATGTACAGT	GGATCAAGGT	8890
	TCTTCTTCAT	TAAGGTTCAA	GTTCTAGATT	GAATAAAGTT	TATGTAACAG	AAACAAAAAT	TTCTTGTACA	8960
55	CACAACCTTG	TCTGGGATTT	GGAGGAAAGT	GTCTCTGAGC	TGGCGGCACA	CTGGTCAAGC	CTGGGTGACA	9030
	GGATACCTCT	GGCCCATGGT	CATGGGGCGC	TGGGCTTGGG	CCTGAGGGTC	ACACAGTGCA	CCATGCCCCG	9100
	CTTCTGTGG	ATAGGATCTG	GGTCTCGGAT	CATGCTGAGG	ACCACAGCTG	CCATGCTGGT	AAAGGGCACC	9170
	ACGTGGCTCA	GAGGGGGCGA	GTTTCCCAGC	CCCAGCTTTT	TTACCGTCTT	CAGTTATTTT	TCCCTAAGAG	9240
	TCTGAGAAGT	GGGGCCGCGC	CTGATGGCCT	TCGTTCGTCT	TCAGCTGGCA	CAGAATTGCA	CAAGCTGATG	9310
	GTAACACATG	AGTACTTATA	ATGAATGAGG	AATGTCTGTA	GCAGTTAACT	GTAAGAGAGT	CGTCTGTTGG	9380
60	AAAGAAATTT	AAGTTTTTCA	TTTTAACCGCT	TTGGAGAATG	TTACTTTTAT	TATGGCTGTG	TAAATTTGTT	9450
	GACATTCAGT	CCCTCGTAGA	CAGATACTAC	GTAAAAGTGT	TAAAGTTAAC	CTTGCTGTGT	ATTTTCCCTT	9520
	ATTTTAGGCT	GCTCCTGCGT	TTGGTGGATG	ATTTCTTGGT	GGTGACACCT	CACCTCACCC	ACGCGAAAAA	9590
	CTTCTCTCAG	TGAGGCCCGT	GCCGTGTGTC	TGTGGGGACC	TCCACAGCCT	GTGGGCTTTG	CATGTGAGCC	9660
65	CCCCGTGTCC	TGCCCTTGGC	ACCGCAGCGT	TGTCTCTGCC	AAGTCTCTCT	TCTCTGCCCG	TGCTGGATCC	9730
	GCAAGAGCAG	AGGCGCTTGG	CCGTGACACC	AGGCTCTGGG	GCGCAGGGGC	ACCTTCGGGA	GGGAGTGGGT	9800
	ACCGTGCAGG	CCCTGGTCTT	GCAGAGACGC	ACCCAGGTTA	CACACGTGGT	GAGTGCAGGC	GAGTGACCTG	9870
	CTCCTGCTGC	TCTTTGGAAG	GTCAAGAGTG	GCGGCTCCTG	GGGCCCCAGT	GAGACCCCCA	GGAGCTGTGC	9940
	ACAGGGCCTG	CAGGGCCGAG	GCGGCAGCCT	CCTCCCCAGG	GTGCACCTGA	GCTTGGCGAG	AGCAGGAGCT	10010
70	GCTGAGTGAG	CTGGCCACAC	CGGTTCTGCT	CGGTGACGTT	CCTGCTGGGG	GTTGTTTGGG	ATCCGTGGGA	10080
	GAATTTGGAT	TTGCTGAGTG	CTGCTGTCTT	GAACACCGGA	GATGGCTAGG	AGTGGGTTTC	AGAGTTGATT	10150
	TTTGTGAATC	AACTTAAAT	CAGGCACAGG	GGAGCTGGCC	TCAGCACAGG	GGATTGTCCA	ATGTGTTCCC	10220
	CCCTCAAGGG	GCCCCACAGA	GCCGTTGGGC	TTGTTTTAAA	GTGCGATTTC	ACGAGGGACG	AGAAACCTTG	10290
	AAAGCTGTAA	AGGGAACCTT	CAGAAAATGT	GGCCGCCAGG	GGTGGTTTCA	GGTGCTTTGC	TGGGCTGTGT	10360
	TTTGTGAATC	CAATTTGGAC	CCGCCCTCCA	AGTCCACCTC	CCAGGTCCAC	CCTCCAGGGC	CGCCCTGGGC	10430
75	TGGGGGTATG	CCTGGCGTTT	CTTGTGCGGC	AGCCCGGAGC	ACAGCAGGCT	GTGCACATTT	AAATCCACTA	10500
	AGATTCACTC	GGGGGGAGCC	CAGGTCCCAA	GCAACTGAGG	GCTCAGGAGT	CCTGAGGCTG	CTGAGGGGAC	10570
	AGAGCAGACG	GGGAACGCTG	CTTCTGTGTG	GCAAGTTCTT	GAGGCTGTGT	GCCAGGGAGG	TGGCTCAGAG	10640
	TGTATGTTTG	GGTCCCACCG	GGGGCAGAAC	TCTGTCTCTG	ATGAGTCCGC	AGCCATGTAA	CAGGAAGGGG	10710

001260-9428560



	TGGCCACAGG	GAGCTGGGAA	TGCACCAGGG	GAGCTGCGCA	GCTGGCCGAG	GTCCACAGGGC	CAGGCCACAG	10780
	GAAGGGCAGG	GGGACGCCCG	GGGCCACAGC	AGAGGCCCGA	GGAAGGGAAG	GGGATGCCCA	GGCCAGAGCA	10850
	GAGGCTACCG	GGCACAGGGG	GGCTCCCTGA	GCTGGGTGAG	CGAGGCTCAT	GACTCGGCGA	GGGAACCTCC	10920
5	TTGACGTGAA	GCTGACGACT	GGTGTGCCCC	AGCTCACAGC	CCAGCCAGGT	CCCGCGCCTG	AGCAGGAACT	10990
	CAGAACCCTC	CCCTTTGTCT	AAAGCACAGC	AGATGCCTTC	AGGGCATCTA	GGAGAAAAAC	GGCAAAGTCG	11060
	TTGAGAAACG	TCTTAAAGA	AGGTGGGATG	GTGGCAATTT	CTTGTCAGAG	TTTTAGTCTG	CCCCGGACCA	11130
	CAGATGAGTC	TATAACGGGA	TTGTGGTGT	GCCATGGGGA	CACATGAGAT	GGACCATCAC	AGAGGCCACT	11200
	GGGGCTGCAC	CTCCCATCTG	AGTCCTGGCT	GTCCCGGGTG	CAGGCCAGGT	TCTTGATATG	TCACCTACCT	11270
10	GTCCTGCCCG	GGAGACAGGG	AAAGCACCCC	GAAGTCTGGA	GCAGGGCTGG	GTCAGGGCTC	CTCAGAGCTC	11340
	CTGCCAGGCC	CAGCACCTCG	CTCCAAATCA	CCACTTCTCT	GGGGTTTTTC	AAAGCATTTA	ACAAGGGTGT	11410
	CAGGTTACCT	CCTGGGTGAC	GGCCCCGCAT	CCTGGGGCTG	ACATTGCCCC	TCTGCCTTAG	GACCCTGGTC	11480
	CGAGGTGTCC	CTGAGTATGG	CTGCGTGGTG	AACTTGCGGA	AGACAGTGGT	GAACCTCCCT	GTAGAAGACG	11550
	AGGCCCTGGG	TGGCACGGCT	TTTGTTCAGA	TGCCGGGCCA	CGGCCATATC	CCCTGGTGGC	GCCTGCTGCT	11620
	GGATACCCCG	ACCCTGGAGG	TGCAGAGCGA	CTACTCCAGG	TGAGGCCACC	TGGCCGGAAG	TGGAGCCTGT	11690
15	SCCCGGCTGG	GGCAGGTGCT	GCTGCAGGGT	CGTTGCTGTC	ACCTTGTGTT	CCGTGTGGGG	CAGCGGACTG	11760
	CCAATCCCAA	AGGGTCAGAG	GCCACAGGGT	GCCCCCTGTC	CCATCTGGGG	CTGAGCAGAA	ATGCATCTTT	11830
	CTGTGGGAGT	GAGGGTGCTC	ACAACGGGAG	CAGTTTTCTG	TGCTATTTTG	GTAAGAGGAA	ATGGTGACCC	11900
	AGACCTGGGT	GCACTGAGGT	GTCTTCAGAA	AGCAGTCTGG	ATCCGAACCC	AAGACGCCCG	GGCCCTGCTG	11970
20	GGCGTGAGTC	TCTCAAACCC	GAACACAGGG	GCCCTGCTGG	GCATGAGTCC	CTCTGAACCC	GAGACCCTGG	12040
	GGCCCTGTCT	GGCGTGAGTC	TCTCCGAACC	CAGAGACTTC	AGGGCCCTTT	TGGGCGTGAG	TCTCTCCGCT	12110
	GTGAGGCCCA	CACCTCAAGG	CTCATCCACA	GTCTACAGGA	TGCCATGAGT	TCATGATCAC	GTGTGACCCA	12180
	TCAGGGGACA	GGGCCATGGT	GTGGGGGGGG	TCTCTACAAA	ATTCTGGGGT	CTTGTTTCCC	CAGAGCCCGA	12250
	GAGCTCAAGG	CCCCGTCTCA	GGCTCAGACA	CAAAATGAAT	GAAGATGGAG	ACAGATGCAG	AAATCTGTGC	12320
25	TGTTTCTTTT	ATGAATAAAA	AGTATCAACA	TCCAGGTCAG	GGCAAGGTGG	CTCACACCTA	TAATCCAGC	12390
	ACTTTGGGAG	GCCGAGGTGG	GTGGATCACT	TGAGGCCAGG	AGTTTGAGGC	CAACCTAACC	AACATAGTGA	12460
	AATTTCAATT	CTACTTAAAA	AATACAAAAA	TTAGCCTGGC	CTGGTGGCAC	ACGCCTGTAG	TCCCCGCTAT	12530
	CGCGGAGGCT	GAGGCAGGAG	AATCATTTGA	ACCCAGGAGG	CAGAGGTTGC	AGTGAGCCGA	GATCACACCA	12600
	CTGCACTCCA	GCCTGGGCAA	CAGAGTGAGA	CTTCATCTTA	AAAAAAGTATC	AAAAAGTATC	GACATTCCAA	12670
30	AACCATAGTG	GACAGGTGTT	TTTTTATTCT	GTCCTTCGAT	AATATTTACT	GGTGCTGTGC	TAGAGCCGGG	12740
	AACTGGGGGT	GCCTTCTCTC	GAAAGGCACA	CCTTCATGGG	AAGAGAAATA	AGTGGTGAAT	GGTTGTTAAA	12810
	CCAGAGGTTT	AAACTGGGGT	CCTGTCGTTT	TGAGTTAACA	GTCCAGATCT	GGACTTTGCC	TCTTTCCAGA	12880
	ATGCTCCCTG	GGGTTTGCTT	CATGGGGGAG	CAGCAGGTGT	GGACACCCCT	GTGATGGGGG	AGCAGCAGGT	12950
35	GCAGACGCCC	TCATGATGGG	GGAGTGGCAG	GTGCCAGAGT	CCTTGTGCAT	GGTGCCACAG	ATGTCCCTGT	13020
	TGCAGCTCCC	TCCCCACAAG	GATGCCGGTC	TCCTGTGCTC	CCCACAGTCC	CTGCTTCCCT	CTCACAGCCT	13090
	TACCTGGTCC	TGGCCTCCAC	TGGCTTTGTC	TGCATGATTT	CCACATTTCC	TGGGGTCCCA	GCACCTCTTC	13160
	GCCTGTCCCA	GGCACCTCTG	CAGTGTCTGC	CATACCAATC	AGCTGTGAAC	TGTCCACTGC	TTATTTGTCT	13230
	CCCCATGAAA	TGTATTTTTT	AGGACAGGCA	CCCCTGGTTC	CAGCCTCTGG	CACAGCATCA	GTGAATGTTA	13300
	TTGAAGGACA	AAGGACAGAC	AAACAAATCA	GGAAATTTGG	TTCTCTCTAA	ACACATTGCA	AAGCCACAGA	13370
40	GGTGTGTGCA	GGATGGGTGG	GCATCAGGTC	ATCAGATGTG	GGTCCAATGC	CAGAATATTG	TGTGCTCCCA	13440
	AAGGCCACTT	GGTCAGAGTG	TGTGCTTGCA	GAGGTGGCTC	TAAAGCTCA	GCAGTGGAGG	CAGTGGTTCT	13510
	CCATCTCAGC	GGTGAATCA	CATCCTCTGT	GTCTGAAAGT	TACAGCAGAG	GCTTGAAGGG	CATCTGGGAG	13580
	AAGAAACACG	GCAAAATGAT	TAAAGAAAGT	GAAAGGTAAG	AAGTGGAAG	ATGGGAATTT	TCTTGTCCAG	13650
	ATTTTAGTCT	CCCAAACAC	AGCTCAGATG	GTAGAATGTG	GTCAGAACTG	ATGGACAGAA	CAATAGAACA	13720
45	AAACGGGAAG	CCTATCTCTC	AGAAACGTGT	GTAAATGTGG	TATGTGGCAC	AGCTGATGGA	AAAGAGAGTG	13790
	TGTGTGTAAT	TTTTTTTTCT	GAGAAACTGT	ACTGGAAGCA	AATAAGTTGT	GTCTTTACAG	CATATACCAG	13860
	AGCAGATTCT	AGGTAGAAGA	GGAGACACAT	GCAAACAACA	CCAGCAACAG	AAATAAAACA	AAAGACTCAA	13930
	AGGGAAGGGA	GGTGAACGTT	CCCTGGTTTG	GTGTTGGGGA	AGGACACACA	GGGAGGCCGA	TGAAACCACT	14000
50	GAGGCAACCG	GCATTGCTTT	CACCTGCAGG	AAACTGCAGT	TGCCTGAGCC	ACAGTGAAAA	TGGCCATTCC	14070
	CTGGAGCGTT	TGTGCACGTG	ATTTATTTAA	GGCGCCCTGT	GAGGTCTGTC	ACATTCTATC	TCTCACTTTG	14140
	TTCTCTTAAC	CACCTGAGAG	GTAGAGGAGG	AAAGGCTCCA	GGGGAGCAGC	CGCCCTTGGT	CACCCAGCTG	14210
	GCAAGGGGCA	TGCATGATTG	CAGCCTGGCC	TCTGTCTCCG	GGGCCCTTGC	TCTGCCCGAG	GACCCACAC	14280
	AAGTCAGACC	CATAGGCTCA	GGGTGAGCCG	GAGCCCAAGG	TCGTGTTGGG	GATGGCTGTG	AAAGAAGAAA	14350
	TGGACGCTGT	ATGCACACTT	GGGAAGGTCC	TACCAGCAGC	GTCAAAGAAA	TGCATGTGAA	ACTGACAGCG	14420
55	AGACCCATCC	CTCAAAGAAA	CGCACGTGAA	ACTGATGGCG	AGACCTGTCC	CCATCCCTCA	TGCTGGCTCC	14490
	TTTTCTGGGC	TTGCCAAGAG	CCAGCATCAG	GTGAGGCAA	GCTGGAAGAA	CTTTTCTGGA	AAGCAGCTTG	14560
	TTTGATGTTA	AGTCCCTACA	ATGTCTGTGT	TCTTCCAGT	AATTTCACTT	CTGAAGTGAC	CAGACATTAT	14630
	CACGGGTCTT	ATTTACCATT	TCCAGTGTTC	CAGGCAGGGG	GACTTGCCAC	AGCAAGTCAC	GAACCTGCC	14700
	AAATACAGGG	CTAAGGAGAT	ATTATGCATC	ACAAAACCTG	CTCTGCCATT	AAACATTTTT	CAAAGAATTT	14770
60	TTGAAGAATG	TTTAATGGCA	CAAAACGTTT	ATTTCAATGT	AGCAGTGTTC	AAAGCTGGAT	GTAAAAGAAC	14840
	ACACCCGACG	AGCCTGCCGT	GAATGTCTAT	TGTGTTTCTC	TTTGGACATG	GACATACATG	GGCAGTGAGT	14910
	GGTGGTGAGG	CCCTGGAGGA	CATCGGTGGG	ATGCCTCCAT	CCTGCCCTTC	TGGAGACACC	ATGTGTGCCA	14980
	CGTGCACTCA	CTGGAGCCCT	GTTTAGCTGG	TGCCACCTGG	CTCTTCCATC	CCTGAGATTG	AAACACAGTG	15050
65	AGATTTCCCA	CGCCCAACTC	AGTGTCTCTC	CACAAAACAA	CTGAGTCACA	CCTGTGTTCA	CTCGAGGGAG	15120
	GCCCGGGAGC	CAGGGCTCCA	CAGTTTATTA	TGTGTTTTTG	GCTGAGTTAT	GTGCAGATCT	CATCAGGGCA	15190
	GATGATGAGT	GCACAAACAC	GGCCGTGCGA	GGTTTGGATA	CACCTCAACAT	CACCTAGCCG	GTCTGCTGCT	15260
	AGTTTTGCTA	TGCAGAGTCT	GGATGGCATG	TAGCATTTGG	AGTCCATGGA	GTGAGACACC	AGCCCCCTCG	15330
	GGCTGCAGCG	CATGCCCCAG	GCAGGACAAG	GAAGCGGGAG	GAAGGCAGGA	GGCTCTTTGG	AGCAAGCTTT	15400
	GCAGGAGGGG	GCTGGGTGTG	GGGCAGGCAC	CTGTGTCTGA	CATTCCCCCC	TGTGTCTCAG	CTATGCCCGG	15470
70	ACCTCCATCA	GAGCCAGTCT	CACCTTCAAC	CGCGGCTTCA	AGGCTGGGAG	GAACATGCGT	CGCAAACTCT	15540
	TTGGGGTCTT	GCGGCTGAAG	TGTCACAGCC	TGTTTCTGGA	TTTGCAGGTG	AGCAGGCTGA	TGGTCAGCAC	15610
	AGAGTTTCAG	GTTCAGGAGG	TGTGTGCGCA	AGTATGTGTG	TGTGTGTGTG	CGCGCGTGCC	TGCAAGGCTG	15680
	ATGGTGACTG	GCTGCACGTA	AGAGTGCACA	TGTACGCGTA	TACACGTGAG	CACATACATG	TGTGCAATGT	15750
	TGTACATGAA	GGCATGGCAG	TGTGTGCACA	GGTGTGCAAG	GGCACAAGTG	TGTGCACATG	CGAATGCACA	15820
	CCTGACATGC	ATGTGTGTTC	GTGCACAGTC	GTGTGGGCGT	TCACGTGAGG	TGCATGCGTG	TGGGCTGTGA	15890
75	GTGTGAGTAG	CATGTGTGCA	CATAACATGT	ATTGAGGGGT	CCTCGTGTTC	ACCCCGCTAG	GCTCTCAGCA	15960
	CCAGTGCCAC	TCCTTACAGG	ATGAGACGGG	GTCCAGGGCC	TTGGTGGGCT	GAGGCTCTGA	AGCTGCAGCC	16030
	CTGAGGGCAT	TGTCCCATCT	GGGCATCCGC	GTCCACTCCC	TCTCTGTGG	GCTTCTGTGT	CCACTCCCCC	16100
	TCTCTGTGG	GCATTTACAT	CCACTCCACT	CCCTCTCTCC	TGTGGGCAAT	CGCGTCCACT	CCCCCTCTCT	16170

	GTGGGCATCT	GCCTCCACCT	CCCCCTCTCTG	TGGGCATTGG	CGTCCACTCC	CTCTCTCTGGT	TCCTTCTCTGT	16240
	CTTGCGCCGAG	CCTCGGGGGC	AGGCAGATGA	CACAGAGTCT	TGACTCGCCC	AGGGTGGTTC	GCAGCTGCGC	16310
	GGTGAGCGGC	AGGCCGGATT	TCACTGGGAA	GAGGGATAGT	TTCTTTGCAA	AATGTTCCCT	TTTCTTGTTCT	16380
5	CATCTGAATG	GATGATAAAG	AAAAAAGTAA	AAACTTAAAA	TCCCAGAGAG	GTTTCTTACCG	TTTCTCACTC	16450
	TTTCTTGCGC	ACTCTAGGTT	AACAGCCTCC	AGACCGTGTG	CACCAACATC	TACAAGATCC	TCCTGCTGCA	16520
	GGCGTACAGG	TGAGCCGCCA	CCAAAGGGTG	CAGGCCGAGC	CTCCAGGGAC	CTCCCGCGCT	CTGCTCACTT	16590
	CTGACCCGGG	GCTTCACTTT	GGAACTCTGT	GGTTTTAGGG	GCAAGGAATG	TCTTACGTTT	TCAGTGGTGC	16660
	TGCTGCCTGT	GCACAGTTCT	GTTCGCGTGG	CTCTGTGCAA	AGCACTGTGT	CTCCATCTCT	GGGTAGTGGT	16730
10	AGGAGCCGGT	TGGCCCCAG	TGTGCTTCTG	TGTGCTTGTG	CAGCTGGCGT	GGGACGTCT	GGAGGCCATC	16800
	CCAGGGCAGC	AGGGGCATGG	GGTAAAGAGA	TGTTTATGGG	GAGTCTTAGC	AGAGAGGAGT	GGGAAGGTGT	16870
	CTGAACAGTA	GATGGGAGAT	CAGATGCCCG	GAGGATTGGG	GGTCTCAGAG	AAGAGGGCCG	AGGTGGGTGC	16940
	AGGTGAGGGT	CGCTGGCCCC	ACCCCGGGGA	AGGTGCAGCA	GAGCTGTGGC	TCCCCACACA	GCCCGGCCAG	17010
	CACCTGTGCT	CTGGGCATGG	CTGTGCTCCT	GGAACTGTTC	CTGTCTGGC	TGGTCAGGGG	GTGCCCTGTC	17080
15	CAAGAATCGA	CAACTTTATC	ACAGAGGGAA	GGGCCAATCT	GTGGAGGCCA	CAGGGCCAGC	TTCTGCCTGG	17150
	ACTCAGGGCA	GGTGTGGCA	CAGCCTCGG	GGCTGTACCA	AAGGGCAGTC	GGGCACCACA	GGCCCGGGCC	17220
	TGCCCTCAA	CAGCGCTCCC	GAGCCACTGG	GAGCTGAATG	CACAGAGGCC	GAGGCCCTCG	CCCCATGAGG	17290
	GCTGAGAAAG	AGTGTGAGCA	TTTGTGTTAC	CCAGGCGCGA	GGCTGCGCGA	ATTACCGTGC	ACACTTGATG	17360
	TGAAATGAGG	TCGTGCTCTA	TCGTGGAAATC	CCAGCAGGGG	CTCACGGGAG	AGTTTTCCAT	TACAAGGTCTG	17430
20	TACCATGAAA	ATGTTTTTTA	ACCCGAGTGC	TTGCGCCTTC	ATGCTCTGGC	AGGGAGGGCA	GAGCCACAGC	17500
	TGCATGTTAT	CGCCTTTTGA	CCAGCTCCAG	AGGGCTTGGGA	CCAGGCTGTC	TCAGTTFCCAG	GGTGCCTCCG	17570
	GCTCAGACCG	CCCTCTCTCT	TGCCTTCTCT	CTCTGCTCTA	AATCTTCCCT	CGTTTGCATC	TCCTTGACGC	17640
	GTGCTTGGGC	CCTCGTGCAG	GCTGCTTGAC	TCCTTTCCGG	AAACCCCTGG	GGTGTGCTGG	ATACAGGTGC	17710
	CACCTGAGAG	TGAGAGGTGC	TGACACTGTG	GTTGACCCCA	GGGTCCAGCT	GGCGTGCTTG	GGGCCTCCTT	17780
25	GGGCCATGAT	GAGGTACAG	GAGTTTTCCC	AGGTGAAACC	TCCTTGGGAA	CTCCCAGGCG	CATGTGACCT	17850
	GCCACCTGCT	CCTCCCATAT	TCAGCTCAGT	CTTGTCTCTA	TTTTCCCAAC	AGGGTCTCTA	GCTCCGAGGA	17920
	GCTCCCCTAG	AGGGCCTGGG	CTCAGGGCAG	GGCGGCTGAG	TTTTCCCAAC	CATGTGGGGA	CCCTTGGGTA	17990
	GTCGTTGAT	TGGGTAGCCC	TGAGGAGGCC	GAGATGCGAT	GGGCCACGGG	CCGTTTCCAA	ACACAGAGTC	18060
	AGGCGAGTGG	AAGGCCCAGG	AATCCCTTCC	CCTCAGGGCA	GGAGTGGGAG	AACGGAGAGC	TGGGCCCGGA	18130
30	TTTCACGGCA	GCCAGGCTGC	AGTGGGCGAG	GCTGTGGTGG	TCCACGCTGGC	GCTGGGGGCG	GGGTCTGATT	18200
	CAAAATCCGT	GGGGCTCGGC	CTTCTCTGCC	CGTGTGCGCC	CGCCTTCCAC	ACGGGCTTGG	GTTGGACGCC	18270
	CCGACCTCTA	GCGAGTGGCT	ATTTCTCCCT	TTGGAAGAGA	GGCCCTCACC	CATGCTAGGT	GGTTCCCTCC	18340
	TGGGTACAGG	GCGTGGCCGT	GTGGCAACCC	CGGGACCTTA	GGCTTATTTA	TTTGTTTAAA	AACATTTCTG	18410
	GCCTGGCTTC	CGTTGTTGCT	AAATGGGGAA	AAGACATCCC	ACCTCAGCAG	AGTTACTGAG	AGGTGAAAC	18480
35	CGGGGTGCTG	GCTTGATGG	TGTGATCTCA	GGTCAATCCA	GAACTGGGCT	AGGAAGTCAG	TGAGACCAGG	18550
	TACATGGGGG	GCTCAGGCAG	TGGGTGAGAT	GAGGTACACG	GGGGGCTCAG	GCAGTGGGTG	AGGCCAGGTA	18620
	CATGGGGGGC	TCAGGCACCT	GGTACAGATG	GGTACACGGG	GGGCTCAGGC	AGAGGGTCAG	ACCAGGTACA	18690
	CGGGGGCTCT	GATCACACGC	ACATATGAGC	ACATGTGCAC	ATGTTGCTGT	TCATGGTAGC	CAGGTCTGTG	18760
	CACACCTGCT	CCAAAGTCCC	AGGAAGCTGA	GAGGCCAAAG	ATGAGGGCTG	ACAGGGCTGG	CCCGGTGGCT	18830
40	CACACCTGTA	GTCCAGCAC	TTTTGGGAGA	CGAGCGCAGA	GGATCCCTTG	AGCCAGGAG	TTTAAGACCA	18900
	GCTTGAGCAA	CATTAGTAGA	CCCCATCTCT	ATGAAAAATA	AAAACAAAAA	TTAGCTGAAC	ATGGTGGTGT	18970
	CGCGCTGTAG	TTCCAATACT	TGGGAGGCTG	AAAGTGGGAG	ATCACTTGAG	CCAGAGGAGT	GGAAAGTCTA	19040
	GTGAGCTGAG	ATTGACACC	TGTACTGCAG	CCTGGGTGAC	AGAGTGAAG	CCCATTCAA	CAACAACAAA	19110
	GAAAGACTGAC	AAATGCAGTT	TCTTGGAAG	AAACATTTAG	TAGGAACCTA	ACCTACACAC	AGAAGCCAA	19180
45	TCGGTGTCTC	GGTGTCACTG	AGATGAGATG	ATGGGTCTCT	ACACATCATC	CCAGACACCA	GGGTTTATGC	19250
	ACCACAGGCG	CGGCTGGCTC	AGAAAGGATG	CGCAGGACGT	TGATATACGA	TGACATCAAG	TGTGCTGAC	19320
	GAAAGGCGAG	ATTATGATA	AGTACCTGCT	GGTACACAAG	GAACAATGGA	TAAACTGGAA	ACCTTAGAGG	19390
	CTTCCCGGCA	ACAGGGGCTA	ATCAGAAGCC	AGCATGGGGG	GCTGGCATCT	AGGATGGAGC	TGCTTCAGCC	19460
	TCCACATGCG	TGTTCATACA	GATGGTGACC	AGAAACCGAC	TGTACTCTTG	CACACACAGA	CACGCAGTCA	19530
50	CTCGCACACA	CAAGCACACA	CACAGACATG	CATGCATGCA	TCCGTGTGTG	TGCACCTGTG	CCCATGAGGA	19600
	AACCTATGCA	TGTGCATTTA	TGCAGCAGCA	CAGGCACCGG	TGGGCCCCAT	CCCCACCTCA	CGAGCACCGT	19670
	CTGATTAGGA	GGCCTTTTCT	CTGACGCTGT	CCGCATCCCT	CTCAGGTTTC	ACGCATGTGT	GCTGCAGCTC	19740
	CCATTTTCATC	AGCAAGTTTG	GAAGAAACCC	ACATTTTTC	TGCGCGTCAT	CTCTGACACG	GCCTCCCTCT	19810
	GCTACTCAGT	CCTGAAAGCC	AAGAAGCCAG	GTATGTGCAG	GTGCTGGGCC	TCAGTGGGAC	CAGTGCTCTG	19880
55	CTGTCTGGTG	TAGTGTGTCA	GGAGACTGAG	TGAACTTGGG	TTTAGGAAGT	TCTTACCCCT	TTTCCGATCA	19950
	GGAGTGGTT	TAACCCAAAC	ACTGTTCAGG	TCGTCTGCCC	GCCCTCTCGT	GGGGTGAGCA	GAGCACCTGA	20020
	TGGAAGGAG	AGGAGCTGTC	TGGGAGCTGC	CATCTTTC	ACCTTGTCTT	GCTCGGGGGA	CGCTGGGGG	20090
	GCCTGGTCTC	TCTGTCTTGC	CCCATGGTGG	GATTTGGGG	GCCTGGCCTC	TCCTGTTTGA	CTCTGTGGTG	20160
	GATTGGGCTG	TCTCCCGTCC	ATGGCACTTA	GGGCCCTTGT	GCAAAACCCAG	GCCAAGGGCT	TAGGAGGAGG	20230
60	CCAGGCCGAG	GCTACCCAC	CCCTCTGGAG	AGCAGAGGCC	CGGTATACCC	ACGACAGAGC	CCCCGCCGCT	20300
	CCTCTGCTTC	CCAGTACCG	TCTCTGCCC	CTGGACACTT	TGTCACGAT	CCGAGAGGTT	TCGTATCCGT	20370
	CTGAAATTCA	AGCCATGTCG	AACCTGCGGT	CCTGAGCTTA	ACAGCTTCTA	CTTTCTGTTT	TTTCTGTGTT	20440
	GTGGAAATTT	CACCTGGAGA	AGCGAAGAA	AACATTTCTG	TCGTTACTCT	TGCGGTGTCT	GGGTGGGGAG	20510
	AGCCAGAGAT	GAGGCCACCC	CGCAGACAGT	CGGGTGTGGG	TGATTTTCCG	GTGTTCTCTG	GGAGGGGAGC	20580
65	TGGGCTGGGC	CTGTGACTCC	TCAGCCTCTG	TTTTCCCCCA	GGGATGTGCG	TGGGGGGCAA	GGGCGCCGCC	20650
	GGCCCTCTGC	CTCCGAGGCC	CTGTGAGTGG	CTGTGCCACT	AAGCATTTCT	CTCAAGCTG	ACTCGACACT	20720
	GTGTCACATA	CGTGCCACTC	CTGGGCTCAC	CTAGGCACGG	CAAGTGTGGG	TGGAGGCCAG	TGCGGGCCCC	20790
	ACCTGCCACG	GGGTCACTCT	TGAACGCCCT	GTGTGGGGCG	AGCAGCCTCA	GATGCTGCTG	AAGTGCAGAC	20860
	GCCTCCGGGC	CTGACCTTGG	GGGCTTGGAG	CCACGCTGGC	AGCCCTATGT	GATTAACAGC	TGGTGTCCCC	20930
70	AGGCCACGGA	GCTGTGGCAG	TGCCCCAACT	TCTTGAACCC	CTGCTTCCCA	TCTCAGGGCG	GATGGCTCCC	21000
	CAGCCTTGGG	AGCCTTCTGA	CCCTTGACCT	GTGTCTCTCT	ACAGCCTCTT	CCCTGGCTGC	TGCCCTGAGC	21070
	TCTTGGGGTG	CTGAGCAAGT	TCTGTCCCCG	CCCCCGCCGT	CCAGCGCTAC	TGGGCTGCCT	GTCTGCTCGC	21140
	CCGGTGGAG	GGGTGTCTGT	CCCTTCACTG	AGGTTCCCA	CAGCCAGGAC	CACAGAGTGC	AGGCCCTGCC	21210
	TGCCCGGCCA	CCACACGTC	CTAGGAGGGT	TGGAGGATGC	CACCTCTGGC	CTCTTCTGGA	ACGGAGTCTG	21280
	ATTTTGGCCC	CGGAGCCGAG	ACCGCAGTGA	GTCCGAAGCT	CCCCGGGACG	ACGCTGACTG	CCCTGGAGGC	21350
75	CGGAGCCAAC	CGGCACTGCG	CTTCAGACTT	CAAGACCATC	CTGGAATGAT	AGCCACCCGC	CCACAGCCAG	21420
	GCCGAGAGCA	GACACCAGCA	GCCCTGTCTC	GCCGGGCTCT	ACGTTCCAGG	GAGGGAGGGG	CGGCCACAC	21490
	CCAGGGCCGC	ACCGCTGGGA	GTCGTGAGCC	TGAGTGAAGT	TTTTGGCCGAG	GCTGTCATGT	CCGGCTGAAG	21560
	GCTGAGTGTG	CGGCTGAGGC	CTGAGCGAGT	GTCGAGCCAA	GGGCTGAGTG	TCCAGCACAC	CTGCCGTCTT	21630

	CACTTCCCCA	CAGGCTGGCG	CTCGGCTCCA	CCCCAGGGCC	AGCTTTTCCT	CACCAGGAGC	CCGGCTTCCA	21700
	CTCCCCACAT	AGGAATAGTC	CATCCCCAGA	TTCCGCCATTG	TTCACCCCTC	GCCCTGCCCT	CCTTTGCCTT	21770
	CCACCCCCAC	CATCCAGGTG	GAGACCCTGA	GAAGGACCCT	GGGAGCTCTG	GGAATTTGGA	GTGACCAAAG	21840
	GTGTGCCCTG	TACACAGGCG	AGGACCCTGC	ACCTGGATGG	GGGTCCCTGT	GGGTCAAATT	GGGGGGAGGT	21910
5	GCTGTGGGAG	TAAAATACTG	AATATATGAG	TTTTTCAGTT	TTGAAAAAAA	TCTCATGTTT	GAATCCTAAT	21980
	GTGCACATGCA	TAGACACCAC	TGTATGCAAT	TACAGAAGCC	TGTGAGTGAA	CGGGGTGGTG	GTCACTGCGG	22050
	GCCCCATGGCC	TGGCTGTGCA	TTTACGGAAG	TCTATGAGTG	AATGGGGTTG	TGGTCAGTGC	GGGCCCATGG	22120
	CCTGGCTGGG	CCTGGGAGGT	TTCTGATGCT	GTGAGGCAGG	AGGGGAAGGA	GGGTAGGGGA	TAGACAGTGG	22190
	GAGCCCCCAC	CCTGGAAGAC	ATAACAGTAA	GTCCAGGCC	GAAGGCGAGC	AGGGATGCTG	GGGGCCAGC	22260
10	TTGGGCGGCG	GGGATGATGG	AGGGCCTGGC	CAGGGTGGCA	GGGATGATGG	GGGCCCCAGC	TGGGGTGGCA	22330
	GGGGTGATGG	GGGGGGCTGG	TCTGGGTGGC	GGGGAAGATG	GGGAAGCCTG	GCTGGGCCCC	CTCCTCCCCT	22400
	GCCTCCACAC	TGCAGCCGTG	GATCCGGATG	TGCTTCCCTG	GTGCACATCC	TCTGGGCCAT	CAGCTTTCAT	22470
	GGAGTGGGG	GGCAGGGGCA	TGACACCATC	CTGTATAAAA	TCCAGGATTC	CTCCTCCTGA	ACGCCCAAC	22540
	TCAGGTTGAA	AGTCACATTC	CGCCTCTGGC	CATTCTCTTA	AGAGTAGACC	AGGATTCTGA	TCTCTGAAGG	22610
15	GTGGGTAGGG	TGGGGCAGTG	GAGGGTGTGG	ACACAGGAGG	CTTCAGGGTG	GGGCTGGTGA	TGCTCTCTCA	22680
	TCCTCTTATC	ATCTCCAGT	CTCATCTCTC	ATCCTCTTAT	CATCTCCAG	TCTCATCTGT	CTTCCTCTTA	22750
	TCTCCAGTC	TCATCTGTCA	TCCTCTTACC	ATCTCCAGT	CTCATCTCTT	ATCCTCTTAT	CTCCTAGTCT	22820
	CATCCAGACT	TACCTCCAG	GGCGGGTGCC	TGGAGCTGGA	AGGCTCTGGA	CATACGTCTC	TCCTCAGGCA	22890
	GAAGGAAGT	GAAGGATTGC	AGAGAACAGG	AGGGGCGGCT	CAGAGGGACG	CAGTCTTGGG	GTGAAGAAAC	22960
20	AGCCCTCTCT	CAGAAGTTGG	CTTGGGCCAC	ACGAAACCGA	GGGCCCTGCG	TGAGTGGCTC	CAGAGCCTTC	23030
	CAGCAGGTCC	CTGGTGGGGC	CTTATGGTAT	GGCCGGGTCC	TACTGAGTGC	ACCTTGGACA	GGGCTTCTGG	23100
	TTTGAGTGCA	GCCCGGACGT	GCCTGGTGTG	GGGGTGGGGG	CTTATGGCCA	CTGGATATGG	CGTCATTTAT	23170
	TGCTGCTGCT	TCAGAGAATG	TCTGAGTGAC	CGAGCCTAAT	GTGTATGGTG	GGCCCAAGTG	CACAGACTGT	23240
25	GTCTGTAATG	CACCTCTGGT	CCTGGAGGCC	CCGTATAGGA	GCTGTGAGGA	AGGAGGGGCT	CTTGGCAGCC	23310
	GGCCTGGGGG	CGCCTTTGCC	CTGCAAACTG	GAAGGGAGCG	GCCCCGGGCG	CCGTGGGCGG	ACGACCTCAA	23380
	GTGAGAGGTT	GGACAGAACA	GGGCGGGGAC	TTCCCAAGGAG	CAGAGGCCGC	TGCTCAGGCA	CACCTGGGTT	23450
	TGAATCAGAG	ACCAACaGTT	CAGGCCATTG	TTCAGCTATC	CATCTTCTAC	AAAGCTCCAG	ATTCTGTGTT	23520
	CTCCGGGTGT	TTTTTGTGTA	AATTTTACTC	AGGATTACTT	ATATTTTTTG	CTAAAGTATT	AGACCCTTAA	23590
	AAAAGGTATT	TGCTTTGATA	TGGCTTAACT	CACTAAGCAC	CTACTTTATT	TGCTGTGTTT	TATTTATTAT	23660
30	TATTTATTAT	ATTAGAGATG	GTGTCTACTC	TGTCACCCAG	GTTGTAGTAG	CAGTGGCACA	GTCTATGCTC	23730
	GCTGTAGCCG	CAAACCCCA	GGCTCAAGTG	ATCCTCCGGC	CTCAGCTTCC	CAGAGTGCTG	GGATTACAGG	23800
	TGTGAGCCAC	TGCCCTTGCC	TGGCACTTTT	AAAAACCACT	ATGTAAGGTC	AGGTCCAGTG	GCTTCCACAC	23870
	CTGTCTATCC	AGTAGTTTGG	GAAGCCGAGG	CAGAAGGATT	GTCTGAGGCC	AGGAGTTTGA	GACCAGCATG	23940
	GGTAACATAG	GGAGACCCCA	TCTCTACAAA	AAATGCAAAA	AGTTATCCGG	GCGTGGGGTC	CAGCATCTGT	24010
35	AGTCCAGCT	GCTCGGGAGG	CTGAGTGCGA	GGATCGCTTG	AGCCCGGGAG	GTCTATGGCTG	CAGTGAGCTG	24080
	TGATTGTACC	ATCGCACTCC	AGCCTGGGCA	ACAGAGTGAG	ACCTGTCTC	AAAAAATAAA	AAAAAATAAG	24150
	AAGGAGAAGG	AGAAGAGAAG	AAGAAGGAAG	AAGGAAAGAG	AAGAAGAAGG	AAGAAGGAAG	AAAGAAGGAG	24220
	AAGGAGGCC	GCTAGGTGCT	AGGTAGACTG	TCAAATCTCA	GAGCAAAATG	AAAAATAACAA	AGTTTAAAG	24290
40	GGAAAGAAAA	ACCCAGCTC	TTTGGACTTC	CTTAGGCCTG	AACTTCATCT	CAAGCAGCTT	CCTTCCACAG	24360
	ACAAGCGTGT	ATGGAGCGAG	TGAGTTCAAA	GCAGAAAGGG	AGGAGAAGCA	GGCAAGGGTG	GAGGTGTGG	24430
	GTGACACCAG	CCAGGACCCC	TGAAAGGGAG	TGGTTGTTTT	CCTGCCTCAG	CCCCACGCTC	CTGCCGGTCC	24500
	TGCACCTGCT	GTAACCGTCG	ATGTTGGTGC	CAGGTGCCCA	CCTGGGAAGG	ATGCTGTGCA	GGGGGCTTGC	24570
	CAAACTTTGG	TGGGTTTCAG	AAGCCCCAGG	CACCTGTGGC	AGGCACAATT	ACAGCCCCTC	CCCAAAGATG	24640
	CCCACGTCTCT	TCTCTGGAA	CCTGTGAATG	TGTCACCCGC	AAGGCAGAGG	CTGGTGAAGG	CTGCAGGTGG	24710
45	AATACGGCT	GCCAGTCAGC	CGATCTTAAG	GTCACTCTGG	ATTATCTGGT	GGGCCTGATA	TGGCCACAAG	24780
	GGTCCCTAGA	AGTGAGAGAG	GGAGGCAGGG	GAGAGTCAGA	GAGGGGACGT	GAGAAGGACC	ACTGGCCACT	24850
	GCTGGCTTTG	AGATGGAGGA	GGGGGTCCCC	AGCCAAGGAA	TGGGGGACGC	CGCTCCATGC	TGGAAAAGCA	24920
	AGCAATCCTC	CCCGGTCCTG	AGGGCACACG	GCCTTCGATA	CGCCTCGATT	TCAGGCCAGT	GGGACCTGTT	24990
	TCAGTCTTCC	GGCCTCCAGA	GCTGTAAGAT	GATGCGTTTG	TGTTAGCCCA	CTAAGCTGCA	GTGATTCTGC	25060
50	ACAGCAGCAA	ATGGAATAGC	AGTACAGGGA	AATGAATACA	GGGACAGTTC	TCAGAGTGAC	TCTCAGCCCA	25130
	CCCCCTGGG							25138

### Example 5

- 55 Comparison of the above-described genomic hTC sequence and the sequence of the hTC cDNA (Fig. 6; corresponding to SEQ ID NO 2) made it possible to elucidate the exon-intron structure of the hTC gene. The genomic organization of the hTC gene is illustrated diagrammatically in Fig. 7. The coding region of the hTC gene is composed of 16 exons which vary in size between 62 bp and 1354 bp (see Table 1).
- 60 Exon 1 contains the translation start codon ATG. The translation stop codon TGA and the 3'-untranslated region lie on exon 16 (Fig. 8). No possible polyadenylation signal (AATAAA) was found either in exon 16 or in the 3195 bp of the following

3'-flanking region. The exon-intron transitions were determined on the basis of the consensus sequence

		5'-Exon			Intron					3'-Exon		
5	Pre-mRNA	A/C	A	G		G	T	A/G	A	...	N C	A G   G
	Frequency (%)	70	60	80		100	100	95	70		80	100 100 60

and listed in Table 1. With the exception of the 5' splice site between exon 15 and  
intron 15, all the exon-intron transitions are in accord with the published (Shapiro  
and Senapathy, 1987) splice consensus sequence. The sizes of the introns are  
between 104 bp and 8616 bp. Since only part of intron 6 was isolated, it is not  
possible to determine the precise length of the hTC gene. Based on the part sequence  
of ~4660 bp, which was obtained from intron 6, the minimum size of the hTERT  
gene is 37 kb.

Introns 1-5 and the 5' region of intron 6, are contained in contig 1:

Intron 1: bp 11493-11596 (SEQ ID NO 4);

Intron 2: bp 12951-21566 (SEQ ID NO 5);

Intron 3: bp 21763-23851 (SEQ ID NO 6);

5 Intron 4: bp 24033-24719 (SEQ ID NO 7);

Intron 5: bp 24900-25393 (SEQ ID NO 8);

5' region of intron 6: bp 25550-26414 (SEQ ID NO 9).

10 The 3' region of intron 6, and introns 7-15, are located in contig 2 at the following positions:

3' region of intron 6: bp 1-3782 (SEQ ID NO 10);

Intron 7: bp 3879-4858 (SEQ ID NO 11);

Intron 8: bp 4945-7429 (SEQ ID NO 12);

Intron 9: bp 7544-9527 (SEQ ID NO 13);

15 Intron 10: bp 9600-11470 (SEQ ID NO 14);

Intron 11: bp 11660-15460 (SEQ ID NO 15);

Intron 12: bp 15588-16467 (SEQ ID NO 16);

Intron 13: bp 16530-19715 (SEQ ID NO 17);

Intron 14: 19841-20621 (SEQ ID NO 18);

20 Intron 15: 20760-21295 (SEQ ID NO 19).

The 3'-untranscribed region is also located in contig 2 at position 21960-25138 (SEQ ID NO 20).

25 The individual sequences of the abovementioned introns are as follows:

00T250" 9423660

**Intron 1 (SEQ ID NO 4)**

GTGGGCCTCCCCGGGTCGGCGTCCGGCTGGGGTTGAGGGCGGCCGGGGGAACCAGCGACATGCGGAGAGCAGCGCAGG  
CGACTCAGGGCGCTTCCCCCGCAG

**5 Intron 2 (SEQ ID NO 5)**

GTGAGGAGGTGGTGGCCGTCGAGGGCCCAGGCCCCAGAGCTGAATGCAGTAGGGGCTCAGAAAAGGGGGCAGGCAGAGCC  
CTGGTCCTCCTGTCTCCATCGTCACGTGGGCACACGTGGCTTTTCGCTCAGGACGTGAGTGGACACGGTGATCTCTGCC  
TCTGCTCTCCCTCCTGTCCAGTTTGCATAAACTTACGAGGTTACCTTCACGTTTTGATGGACACGCGGTTTCCAGGCGC  
CGAGGCCAGAGCAGTGAACAGAGGAGGCTGGGCGCGGCAGTGGAGCCGGGTGCCGGAATGGGGAGAAGTGTCTGGAAG  
10 CACAGACGCTCTGGCGAGGGTGCCGTCAGGTTACCTATAATCCTCTTCGCAATTTCAAGGTGGGAATGAGAGGTGGGGA  
CGAGAACCCCTCTTCTGGGGTGGGAGGTAAGGGTTTTGCAGGTGCACGTGGTCAGCCAATATGCAGGTTTGTGTTTA  
AGATTTAATTGTGTGTTGACGGCCAGGTGCGGTGGCTCACGCCGGAATCCCAGCACTTTGGGAAGCTGAGGCAGGTGGA  
TCACCTGAGGTGAGGAGTTTGAGACCAGCCTGACCAACATGGTGAAACCCTATCTGTACTAAAAATACAAAATTAGCTG  
GGCATGGTGGTGTGTGCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCACTTGAACCCAGGAGGCGGAGGC  
15 TGCAGTGAGCTGAGATTGTGCCATTGTACTCCAGCCTGGGCGACAAGAGTGAACTCTGTCTTTAAAAAAAAAAGTGT  
CGTTGATTGTGCCAGGACAGGGTAGAGGGAGGGAGATAAGACTGTTCTCCAGCACAGATCCTGGTCCCATCTTTAGGTAT  
GAAGAGGGCCACATGGGAGCAGAGGACAGCAGATGGCTCCACCTGCTGAGGAAGGGACAGTGTGTTGGGTGTTTCAGGGG  
ATGGTGCTGCTGGGCCCTGCCGTGTCCCAACCTGTTTTCTGGATTTGATGTTGAGGAACCTCCGCTCCAGCCCCCTTT  
TGGCTCCCAGTGCTCCAGGCCCTACCGTGGCAGCTAGAAGAAGTCCCGATTTACCCCCCTCCCCACAACTCCCAAGAC  
20 ATGTAAGACTTCCGGCCATGCAGACAAGGAGGGTGACCTTCTTGGGGCTCTTTTTTTCTTTTTTTCTTTTTATGGTGGC  
AAAAGTCATATAACATGAGATTGGCACTCCTAACACCGTTTTCTGTGTACAGTGCAGAATTGCTAACTCGGCGGTGTTTA  
CAGCAGGTTGCTTGAAATGCTGCGTCTTGCGTACTGGAAGTCCCTACCCATCGAACGGCAGCTGCCTCACACCTGCTGC  
GGCTCAGGTGGACCACGCCGAGTCAGATAAGCGTCATGCAACCCAGTTTTGCTTTTTGTGCTCCAGCTTCCTTCGTTGAG  
GAGAGTTTGAGTTCTCTGATCAGGACTCTGCCTGTCTGCTGTTCTCTGACTTCAGATGAGGTACAATCTGCCCTGG  
25 CTTATGCAGGGAGTGAGGCGTGGTCCCCGGGTGTCCCTGTACGTCAGGGTGAGTGAGGCGTTGCCCCCAGGTGTCCCT  
GTCACGTGTAGGGTGAGTGAGGCGCGGCCCCGGGTGTCCCTGTCCCGTGCAGCGTGATTGAGGTGTGGCCCCCGGGTGT  
CCCTGTACGTCGTAGGGTGAGTGAGGCGCCATCCCCGGGTGTCCCTGTACGTCGTAGGGTGAGTGAGGCGTGGTCCCCGG  
GTGTCCCTGTCCCGTGCAGGGTGAGTGAGGCACTGTCCCCGGGTGTCCCTGTACGTCGTAGGGTGAGTGAGGCGCGGTCC  
CCGGGTGTCCCTCTCAGGTGTAGGGTGAGTGAGGCGCGGCCCCAGGGTGTCCCTGTACGTCGTAGGGTGAGTGAGGCACC  
30 GTCCCTGGGTGTCCCTCCCAGGTATAGGGTGAGTGAGGCACTGTCCCCGGGTGTCCCTGTACGTCGTAGGGTGAGTGAGG  
CGCGGCCCCCGGTGTCCCTCTCAGGTGCAGGGTGAGTGAGGCGCTGTCCCTGGGTGTCCCTGTCTCGTGTAGGGTGAGT  
GAGGCTCTGTCCCCAGGTGTCTTGGCGTTTGCTCACTTGAGCTTGCTCCTGAATGTTTGCTCTTTCTATAGCCACAGCT  
GCGCCGGTTGCCATTGCCTGGGTAGATGGTGCAGGCGCAGTGCTGGTCCCCAAGCCTATCTTTTCTGATGCTCGGCTCT  
TCTTGGTCACCTCTCCGTTCCATTTTGCTACGGGGACACGGGACTGCAGGCTCTCGCCTCCCGCGTGCCAGGCACTGCAG  
35 CCACAGCTTCAGGTCCGCTTGCCCTCTGTTGGGCCTGGCTTGCTCACCACGTGCCCGCCACATGCATGCTGCCAATACTCC  
TCTCCCAGCTTGTCTCATGCCGAGGCTGGACTCTGGGCTGCCTGTGTCTGTGCCACGTGTGTGCGAGACATCCCAGAA  
AGGGTTCTCTGTGCCCTGAAGGAAAGCAAGTCACCCACGCCCCCTCACTTGTCTGTTTTCTCCCAAGCTGCCCTCTGC  
TTGGCCCCCTTGGGTGGGTGGCAACGCTTGTCACCTTATCTTGGGCACCTGCCGCTCATTGCTTAGGCTGGGCTCTGCCT  
CCAGTCGCCCCCTCACATGGATTGACGTCCAGCCACAGGTGGAGTGTCTGTCTGTCTCCTGCTCTGAGACCCACGTG  
40 GAGGGCCGGTGTCTCCGCCAGCCTTCGTGCACTTCCCTCTTGGGTCTTAGTTTTGAATTTCACTGATTTACCTCTGACG  
TTTCTATCTCTCCATTGTATGCTTTTTCTTGGTTTTATTCTTTTCATTCCTTTCTAGCTTCTTAGTTAGTCATGCTTTT  
CCTCTAAGTGCTGCCTTACCTGCACCTGTGTTTTGATGTGAAGTAATCTCAACATCAGCCACTTTCAAGTGTCTTTAAA  
ATACTTCAAAGTGTTAATACTTCTTTTAAGTATTCTTATCTGTGATTTTTCTTTGTGCACGCTGTGTTTTGACGTGA  
AATCATTTTGATATCAGTGACTTTTAAGTATTCTTTAGCTTATCTGTGATTTCTTTGAGCAGTGAGTTATTTGAACACT  
45 GTTTATGTTCAAGATATGTAGAGTATCAAGATACGTAGAGTATTTTAAGTTATCATTTTATTATTGATTTCTAACTCAGT  
TGTGTAGTGGTCTGTAATAACCAATTATTTGAAGTTTGGGAGCCTTGCTTTGTGATCTAGTGTGTGCATGGTTTCCAG  
AACTGTCCATTGTAATTTGACATCCTGTCAATAGTGGGCATGCATGTTCACTATATCCAGCTTATTAAGGTCCAGTGCA

45

5 GAGGTATGGAGTCCGGATGATGCAGGTCCGGGGTGAGGTTGCCAGGCCCTGCTGTGAGCTGGATGTGCTGTATCCGGATG  
GTGCAGTCCGGGGTGAGGTCGCCAGGCCCTGCTGTGAGCTGGATGTGCTGTATCCGGATGGTGCAGGTCTGGGGTGAGGT  
CACCAGGCCCTGCGGTGAGCTGGTTGTGCGGTGTCCGGTTGCTGCAGGTCCGGGGTGAGTTCGCCAGGCCCTCGGTGAGC  
TGGATGTGCGGTGTCCCCGTGTCCGGATGGTGCAGGTCCAGGGTGAGGTGCGTAGGCCCTTGGTGGGCTGGATGTGCCGT  
10 GTCCGGATGGTGCAGGTCTGGGGTGAGGTGCCAGGCCCTTGGTGAGCTGGATGTGCGGTGTCTGCATGGTGCAGGTCTG  
GGGTGAGGTGCCAGGCCCTTGGTGGGCTGGATGTGTGGTGTCCGGATGGTGCAGGTCCGGCGTGAGGTGCCAGGCCCT  
GCTGTGAGCTGGATGTGCGGTGTCTGGATGGTGCAGGTCCGGGGTGAGGTAGCCAAGGCCCTTCGGTGAGCTGGATGTGGG  
GTGTCCGGATGGTGCAGGTCCGGGGTGAGGTGCCAGGCCCTGCGGTTAGCTGGATATGCGGTGTCCGGATGGTGCAGGT  
CCGGGGTGAGGTACCCAGGCCCTGCGGTTAGCTGGATGTGCGGTGTCTGGATGGTGCAGGTCCGGGGTGAGGTGCCAGG  
15 CCCTGCTGTGAGCTGGATGTGCTGTATCCGGATGGTGCAGGTCCGGGGTGAGGTGCCAGGCCCTGCAGTGAGCTGGATG  
TGCTGTATCCGGATGGTGCAGGTCTGGCGTGAGGTGCCAGGCCCTGCGGTTAGCTGGATATGCGGTGTCCGGATGGTGCA  
GGTCCGGGGTGAGGTACCCAGGCCCTGCGGTTAGCTGGATGTGCGGTGTCCGGATGGTGCAGGTCTGGGGTGAGGTGCC  
AGGCCCTGCTGTGAGCTGGATGTGCTGTATCCGGATGGTGCAGGTCCGGGGTGAGGTGCCAGGCCCTGCGGTGAGCTGG  
ATGTGCTGTATCCGGATGGTGCAGGTCTGGCGTGAGGTGCCAGGCCCTGCGGTGAGCTGGATGTGCAGTGTACGGATGG  
20 TGCAGGTCCGGGGTGAGGTGCCAGGCCCTGCGGTGGGCTGTATGTGTGTTGTCTGGATGGTGCAGGTCCGGGGTGAGTT  
CGCCAGGCCCTGCGGTGAGCTGGATGTGTGGTGTCTGGATGCTGCAGGTCCGGGGTGAGTTCGCCAGGCCCTCGGTGAGC  
TGGATATGCGGTGTCCCCGTGTCCGAATGGTGCAGGTCCAGGGTGAGGTGCCAGGCCCTTGGTGGGCTGGATGTGCCGT  
GTCCGGATGGTGCAGGTCTGGGGTGAGGTGCCAGGCCCTTGGTGAGCTGGATGTGCGGTGTCCGGATGGTGCAGGTCCG  
GGGTGAGGTACCCAGGCCCTCGGTGATCTGGATGTGGCATGTCTTCTCGTTTAAAG

**Intron 3 (SEQ ID NO 6)**

GTACTSTATCCCCACGCCAGGCCCTCTGCTTCTCGAAGTCCTGGAACACCAGCCCCGGCCTCAGCATGCGCCTGTCTCCACT  
TGCCTGTGCTTCCCTGGCTGTGCAGCTCTGGGCTGGGAGCCAGGGGCCCCGTACAGGCCTGGTCCAAGTGGATTCTGTG  
CAAGGCTCTGACTGCCTGGAGCTCACGTTCTCTTACTTGTAAATCAGGAGTTTGTGCCAAGTGGTCTCTAGGGTTTGTA  
25 AAGCAGAAGGGATTTAAATTAGATGGAAACACTACCACTAGCCTCCTTGCCCTTCCCTGGGATGTGGGTCTGATTCTCTC  
TCTCTTTTTTTTTTCTTTTTTGTAGATGGAGTCTCACTCTGTTGCCAGGCTGGAGTGCAGTGGCATAATCTTGGCTCACT  
GCAACCTCCACCTCCTGGGTTTAAGCGATTACCCAGCCTCAGCCTCCTAAGTAGCTGGGATTACAGGCACCTGCCACCAC  
GCCTGGCTAATTTTTGTACTTTTAGGAGAGACGGGTTTACCATGTTGGCCAGGCTGGTCTCGAAGTCACTGACCTCAGG  
TGATCCACCCACCTTGGCCTCCCAAAGTGCTGGGTTTACAGGCTAAGCCACCGTGCCAGCCCCGATTCTCTTTAATT  
30 CATGCTGTTCTGTATGAATCTTCAATCTATTGGATTTAGGTCATGAGAGGATAAAATCCACCCACTTGGCGACTCACTG  
CAGGGAGCACCTGTGCAGGGAGCACCTGGGGATAGGAGAGTTCCACCATGAGCTAAGTCTAGGTGGCTGCATTTGAATG  
GCTGTGAGATTTGTCTGCAATGTTCCGCTGATGAGAGTGTGAGATTGTGACAGATTCAAGCTGGATTTGCATCAGTGAG  
GGACGGGAGCGCTGGTCTGGGAGATGCCAGCCTGGCTGAGCCCAGGCCATGGTATTAGCTTCTCCGTGTCCGCCCAGGC  
TGACTGTGGAGGGCTTTAGTCAGAAGATCAGGGCTTCCCCAGCTCCCCTGCACACTCGAGTCCCTGGGGGGCCTTGTGAC  
35 ACCCATGCCCCAAATCAGGATGTCTGCAGAGGGAGCTGGCAGCAGACCTCGTCAGAGGTAACACAGCCTCTGGGCTGGG  
GACCCCGACGTGGTGTGGGGCCATTTCTTGCATCTGGGGGAGGGTCAGGGCTTTCCTGTGGGAACAAGTTAATACAC  
AATGCACCTTACTTAGACTTTACACGTATTTAATGGTGTGCGACCCAAACATGGTCATTTGACCAGTATTTTGGAAAGAAT  
TTAATTGGGGTGACCGGAAGGAGCAGACAGACGTGGTGGTCCCCAAGATGCTCCTTGTCACTACTGGGACTGTTGTTCTG  
CCTGGGGGGCCTTGGAGGCCCTCCTCCCTGGACAGGGTACCGTGCTTTTCTACTCTGCTGGGCCTGCGGCCTGCGGTG  
40 AGGGCACCAGCTCCGAGACCCCGCGGCCCAAGTGTCCACGGAGTGCCAGGCTGTACGCCACAGATGCCAGGTCCAGGT  
GTGGCCGCTCCAGCCCCCGTCCCCCATGGTGGTTTTGGGGGAAAAGGCCAAGGGCAGAGGTGTGAGGAGACTGGTGGG  
CTCATGAGAGCTGATTCTGCTCCTTGGCTGAGCTGCCCTGAGCAGCCTCTCCCGCCCTCTCCATCTGAAGGGATGTGGCT  
CTTTCTACCTGGGGGTCTGCTGGGGCCAGCCTTGGGCTACCCAGTGGCTGTACCAGAGGGACAGGCATCCTGTGTGG  
AGGGGCATGGGTTCACGTGGCCCCAGATGCAGCCTGGGACCAGGCTCCCTGGTGTGATGGTGGGACAGTCACCCCTGGGG  
45 GTTGACCGCCGACTGGGCTCCCCAGGGTTGACTATAGGACCAGGTGTCCAGGTGCCCTGCAAGTAGAGGGGCTCTCAG  
AGGCGTCTGGCTGCGCATGGGTGGACGTGGCCCCGGGCATGGCCTTCAGCGTGTGCTGCCGTGGGTGCCCTGAGCCCTCAC  
TGAGTCGGTGGGGCTTGTGGCTTCCCGTGAGCTTCCCCCTAGTCTGTTGTCTGGCTGAGCAAGCCTCCTGAGGGGCTCT  
CTATTGCAG



## 5

10

Intron 5 (SEQ ID NO 8)

15

20

5'-region intron 6 (SEQ ID NO 9)

25

30

3'-region intron 6 (SEQ ID NO 10)

35

40

TGTGGGATTGGTTTTTCATGTGTGGGATAGGTGGGGATCTGTGGGATTGGTTTTTATGAGTGGGGTAACACAGAGTTC AAG  
GCGAGCTTTCTTCCGTAGTGGGTCTGCAGGTGCTCCAACAGCTTTATTGAGGAGACCATATCTTCCCTTTGAACATGGT  
CGGGTTTATAGTAAGTCAGGGGTGTGGAGGCCCTCCCCTGGGCTCCCTGTTCTGTTTCTTCCACTCTGGGGTCGTGTGGT  
CCTGCTGTGGTGTGTGGCCGGTGGGCAGGGCTTCAGGCCCTCCTTGTTTCATTGGCCTGGATGTGGCCCTGGCTACGCT  
CCGTCTTTGGAATCCCCTGCGAGTTGGAGGCTTTCTTTCTTTCTTTTTTTCTTTCTTTTTTTTTTTTTTTTGATAACAGA  
GTCTCGCTCTTTTTTGCCCAGGCTGGAGTGTTTGGCGTGATCTTGCTCACTGCAACCTGTGCTTCCTGAGTTC AAGCA  
ATTCTCTTGCCCTCAGCCTCCCAAGTAGCTGGAATTATAGGCGCCACCACCATGCTGACTAATTTTTGTAATTTTAGTAG  
AGACGAGGTTTCTCCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCCTCCACCTCGGCCTCCCAAAGT  
GCTGGGATGACAGGTGTGAACCGCCGCGCCCGGCGGAGACTCGCTTCCTGCAGCTTCCGTGAGATCTGCAGCGATAGCTG  
CCTGCAGCCTTGGTGCTGACAACCTCCGTTTTCTTTCTCCAGGTCTCGCTAGGGGTCTTTCCATTTTCATGACTCTCTTC

CAGAAGAGTTTCACGTGTGCTGATTTCCCGGCTGTTTCCTGCGTAATTGGTGTCTGCTGTTTATCGATGGCCTCCTTCCA  
 TTTCCCTTAGGCTTTGTTTATTGTTGTTTTTCCGGCTCCTTGAAGGAAAAGTTTCGATTATGGATGTTTGAACCTTCTTT  
 TCTAAACAAGCATCTGAAGTTGCCGTTTTCCCTCTAAAGCAGGGATCCCGAGGCCCTGGCTGTGGAGTGGCACCAGGTCT  
 5 GGGGCCCTGTTAGGAACCCGGCGCACAGCGGGAGGCTAGGTGGGGTGTGGGGAGCCAGCGTTCCCGCCTGAGCCCCGCCCC  
 TCTCAGATCAGCAGTGGCATGCGGTGCTCAGAGGCGCACACACCCTACTGAGAACTGTGCGTGAGAGGGGTCTAGATTCT  
 GTGCTCCTTATGGGAATCTAATGCCTGATGATCTGAGGTGGAACCGTTTGCTCCCAAACCATCCCCCTCCCCACTGCTG  
 TCCTGTGGAAAAATCGTCTTCCACGAAACAGTCCCTGGTACCACAATGGTTGGGGACCCTGTGCTAAAGACCTGCTTCA  
 GCAGCCTCTCGTCAGTGTGATATATTGGCTTTTCTGTGTTGAGTCCAGAATAATTACGGATTTCTGTGATGCTTTCCGC  
 CGACCTCAGACCCATGGGCTATTTGTGGCGTGTGCTGCTCCTGGGTGGGAAGGTGCAGGCCCATGTACCTTCT  
 10 GTTACTGCCCTTCCAGGTTGGTTCTCAGGGTTGAATCGTACTCGATGTGGTTTTAGCCCACGGCCCTGCCGCCAGCTCCTG  
 GGGGCTGGGGAACATGCTGAAGCACAGAGTCACCGTGCGCTCTTTGATGCCTCACAAGCTCGAGGCCCTCTGTGTCCG  
 TGTTAGTGTGTGTACGTGCCTGCTCACATCCTGTCTTGGGACGCAGGGGCTTAGCAGGTCCCGTAGTAAATGACAAGC  
 GTCCTGGGGGAGTCTGCAGAATAGGAGGTGGGGTGCCGGTCTCTCTCCCGCTCTTCAGACTCTTCTCCTGCCTGTGCT  
 GTGGCTGCACCTGCATCCCTGCAATCCCTCCAGCACTGGGCTGGAGAGGCCCGGGAGCTCGAGTGCCACTTGTGCCACGT  
 15 GACTGTGGATGGCAGTGGTCCAGGGGGTCTGATGTGTGGTGAATGTGGATGGCGGTGGTTCACAGGGGTCTGATGTGTG  
 GTGACTGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGATGTGTGGTGAATGTGG  
 ATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTG  
 GGGTCTGATGTGGTGAATGTGGATGGCAGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATG  
 TGGTGAATGTGGATGGCAGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAAT  
 20 GTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGG  
 CGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGTGAATCGGTCA  
 CAGGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGTGAATCGGTCA  
 GGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTGGTCCCAGGGG  
 TCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCT  
 25 GATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGT  
 GACTGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAAT  
 GGCGGTGGTCCCAGGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCAG  
 TCGTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGG  
 TCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGT  
 30 GGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGTGAATCGGTCAAGGGGTCTGATGTGGT  
 GACTGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAAT  
 GGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCAG  
 GTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGAT  
 35 GTGGTGAATGTGGATGGTGAATCGGTCAAGGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGT  
 ACTTTGGCTCCTCGGCCCCCGGCCCCCGTTTCCCAAACAGAAGCTTCCAGGCGCTCTGTTGGCTTCATCCCAGCATCG  
 GGCTTGGCCGACAGTCCACACGTCCTGATCGGAAGAAACAAGTGCCAGCTCTGGCCGGGACAGGCCACATTTGTGGCTC  
 ATGCCCTCTCCTCTGCCGGCAG

#### 40 Intron 7 (SEQ ID NO 11)

GTCTGGGCACTGCCCTGCAGGGTTGGGCACGGACTCCCAGCAGTGGGTCTCCCCTGGGCAATCACTGGGCTCATGACCG  
 GACAGACTCTTGGCCCTGSGGGGACAGTGGGGGGAATGAGCTGTGATGGGGGATGATGAGCTGTGTGCCCTTGGCGAAATC  
 TGAGCTGGGCCATGCCAGGCTGCGACAGCTGCTGCATTACAGGCACCTGCTCACGTTTACTGCGCGGCCCTCTCTCCAGTT  
 CCGCAGTGCCTTTGTTCAATGATTTGCTAAATGTCTCTGCCAGTTTTGATCTTGAGGCCAAAGGAAAGGTGTCCCCCT  
 45 CCTTTAGGAGGGCAGGCCATGTTTGAACCGTGTCTGCCAGCTGGCCCTCAGTGTGGGTCTGAGGCCAAAGGAAACG  
 TGTCCCCCTCTTAGGAGTACGGGGCCGTGTTTGAACACGCCCGCTGAGCGGGCCTCTCAGTGTGGGTCTGTCCAGT

GGCCCTGTGGCCCTTTGCAGATGTGGTCTGTCCACGTGGCCCTGTGGCTCTTTGCAGATGCCTGTTAGCACTTGCTCGGC  
TCTAGGGGACAGTCGTGTCCACCGCATGAGGCTCAGAGACCTCTGGGCGAATTTCTTGGCTCCCAGGGTGGGGTGGAG  
GTGGCCTGGGCTGCTGGGACCCAGACCCTGTGCCCGGCAGCTGGGCAGCAACTCCTGGATCACATATGCCATCCGGGCCA  
CGGTGGGCTGTGTGGGTGTGAGCCCAGCTGGACCCACAGGTGGCCAGAGGAGACGTTCTGTGTACACACTCTGCCTAA  
5 GCCCATGTGTCTGCAGAGACTCGGCCCGCCAGCCACGATGGCCCTGCATTCCAGCCCAGCCCCGCACTTCATCACA  
AACACTGACCCCAAAGGGACGGAGGGTCTTGCCACGTGGTCTGCCTGTCTCAGCACCCACGGGCTCACTCCCATGTG  
TCTCCCGTCTGCTTTCGCAG

**Intron 8 (SEQ ID NO 12)**

10 GTGAGTCAGGTGGCCAGGTGCCATTGCCCTGCGGGTGGCTGGGCGGGCTGGCAGGGCTTCTGCTCACCTCTCTCCTGCCC  
CTTCCCCACTGNCCTTCTGCCCCGGGGCCACCAGAGTCTCCTTTTCTGGCCCCCGCCCCCTCCGGCTCCTGGGCTGCAGGC  
TCCCGAGGCCCCGAAACATGGCTCGGCTTGCGGCAGCCGAGCGGAGCAGGTGCCACACGAGGCCTGGAAATGGCAAGC  
GGGGTGTGGAGTTGCTCCTGCGTGGAGGACGAGGGGCGGGGGTGTGTCTGGGTGAGGTGTGCGCCGAGCGTTTGAGCCT  
GCAGCTTGTGAGTCCAAGTTACTACTGACGCTGGACACCCGGCTCTCACACGCTTGTATCTCTCTCTCCCCATACAAAA  
15 GGATTTTATCCGATTCTCATTCTGTCCCTGTGCTGTGACCCCCGCGAGGGCGGGGCTCTTCTCTCTGTGACTAGATTT  
CCCCTCTGGAAAGTGCAGGGTTGACCGTGTAGTTTGCTCCTCTCGGGGGCCTGTGGTGGCCATGGGGCAGGCGGCCTGG  
GAGAGCTGCCGTACACAGCCACTGGGTGAGCCACACTCACGGTGGTAGAGCCACAGTGCCTGGTGCCACATCACGTCCT  
CTGGATTTTAAAGTAAACACACACCTCCCGGCAGGCATCTGCCTGCGACCTGTGTGTGCCTGGGAGAGTGGTAGCAC  
GGAGGAAATTCGTGCACACTCAAGGTATCAGCAAGGTATCCGCAGTCAGGTGGAACGTGGAGGCCCTCTCTCTGGGATC  
20 GTCTCCAGCGGATAAAGGACTGTGCACAGCTTCGGAAGCTTTTATTTAAAAATATAACTATTAATTATTGCATTATAAGT  
AATCACTAATGGTATCAGCAATTATAATATTTATTAAGTATAATAGAAATATTAAGTAGTACACACGTTCTGGAAAAA  
CACAAATTGCACATGGCAGCAGAGTGAATTTTGGCCGAGGGACACGTGTGCACATGTGTGTAAGCGGCCCCAGGCCAC  
AGAATTCGCTGACAAAGTCACTCCCCAGAGAAGCCACCACGGGCCTCCTTCGTGGTGTGAATTTTATTAAGATGGATC  
AAGTCACGTACCGTCCACGTGTGGCAGGGCTTTGGGGAATGTGAGGTGATGACTGCGTCTCATGCCCTGACAGACAGGA  
25 GGTGACTGTGTCTGTCTGTCCCTAGGACACGGACAGGCCCGAAGCTCTAGTCCCCATCGTGGTCCAGTTTGGCCTCTGA  
ATAAAAACGTCTTCAAAACCTGTTGCCCCAAAACTAAGAACAGAGAGAGTTTCCCATCCCATGTGCTCACAGGGGCGTA  
TCTGCTTGCGTTGACTCGCTGGGCTGGCCGGACTCCTAGAGTTGGTGCCTGTGCTTCTGTGCAAAAAGTGCAGTCCTCTT  
GCCCATCACTGTGATATCTGCACCAGCAAGGAAAGCCTCTTTTCTTTTCTTTTCTTTTCTTTTGGAGACGGAACGTCA  
CTGTTGTCTGCCTGGGCTTGAGTGCAGTGGCGCGATCTCAACTCACTGCAACCTCCGCCTCCCGGGTCCAGCATTCTC  
30 CTGCCTCAGCCTCCCGAGCAGCTGAGATTACAGGCACCCACCCCTGCGCCTGGCTAATTTTTGTATTTTGTAGTAGAG  
GGGTTTTTGCCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGTG  
GGATTACAGGTGTGAGCCATCACGCCAGCCGAAAGCCTCTTTTAAAGGTGACCACCTATAGCGCTTCCGAAAATAAC  
AGGTCTTGTTTTTGCAGTAGGCTGCAAGCGTCTCTTAGCAACAGGAGTGGCGTCTGTGGGCTCTGGGGATGGCTGAGGG  
TCGCGTGGCAGCCATGCCTTCTGTGTGCACCTTTAGGTTCCACGGGGCTATTCTGCTCTCACTGTTTGTCTGAAAACGCA  
35 CCTTGGCATCCTTGTTTGGAGAGTTTCTGCTTCTCGTTGGTCACTGTGAACTAGGGGCAAGGTTGTATCCGTTGGCGC  
GCAGCGGCTACATGTAGGCTCATGAGTCTTTCACCGTGGACAAATTCCTTGAAAAAAGGAGTCCGGTTAAGCAT  
TCATTCCGGGTCAAGTGTCTGTTCTGTGAATAAACTCTAAGATTTAAGAAACCTTAATGAAAGAAAACCTTGATGATTC  
AGAGCAAGGATGTGGTACACCTGTGGCTGGATCTGTTTCAGCCGCCCCAGTGCATGGTGAGAGTGGGGAGCAGGGATTG  
TTTGTTAGAGGTCTCATCTGGTATGTTTCTGAGGTGTTTGCCGGCTGAATGGTAGACGTGTCGTTTGTGTGTATGAGGT  
40 TCTGTGTCTGTGTGGCTCGGTTTGAGTGTACGCATGTCCAGCACATGCCCTGCCCGTCTCTACCTGTGTCTTCCCGC  
CCAG

**Intron 9 (SEQ ID NO 13)**

45 GTGAGGCCTCCTCTTCCCCAGGGGGCTTGGGTGGGGTTGATTTGCTTTTGATGCATTCACTGTTAATATTCCTGGTGC  
TCTGGAGACCATGACTGCTCTGTCTTGAGGAACCAGACAAGGTTGCAGCCCTTCTTGGTATGAAGCCGCACGGGAGGGG

TTGCACAGCCTGAGGACTGCGGGCTCCACGCAGGCTCTGTCCAGCGGCCATGTCCAGAGGCCTCAGGGCTCAGCAGGCGG  
GAGGGCCGCTGCCCTGCATGATGAGCATGTGAATCAACACCGAGGAAGCACACCAGCTTCTGTACGTACCCAGGTTC  
CGTTAGGGTCCCTTGGGGAGATGGGGCTGGTGCAGCCTGAGGCCCCACATCTCCAGCAGGCCCTCGACAGGTGGCCTGGA  
CTGGGCGCCTCTTCAGCCCATTGCCCATCCCACTTGCATGGGGTCTACACCCAAGGACGCACACCTAAATATCGTGCC  
5 AACCTAATGTGGTTCAACTCAGCTGGCTTTTATTGACAGCAGTTACTTTTTTTTTTTAATACTTTAAGTTCTAGGGTAC  
ATGTGCACGACGTGCAGGTTAGTTACATATGTATACATGTGCCATGTTGGTGTGCTGCACCCATTAACATCATCTTTACA  
TTAGGTATATCTCCTAATGCTATCCCTCCCCACTCCCCCATCCCATGACAGGCCCTGGTGTGTGATGTTCCCCACCTG  
TGTCCAAGTGTCTCATTGTTTCTGAGTTCACCTGTGAGTGAGAATGTGGTGTGTTGGTCTTTCTTCTTCAATAGTTT  
GCTCAGAGTGATGGTTTCCAGCTTCGTCCATGTCCCTACAAAGGACATGAACTCATCTTTTTTATGACTGCATAGTATT  
10 CCGTGGTGTATATGTGCCACATTTTCTTAATCCAGTCTATCATCGATGGACATTTGGGTGGTTGCAAGTCTTTGCTACT  
GTGAATAGTGCCGCAATAAACATACGTGTGCATGTGTCTTTATAGCAGCATGATTTATAATCCTTTGGGTATATACCCAG  
TAATGGGATGGCTGGGTCAAATGGTATTTCTAGTTCTAGATCCTTGAGGAATCACCACACTGTCTTCCACAATGGTTGAA  
CTAGTTTACACTCCCACCAACAGTGTAAGAGTGTCTGGTGTGAGAGGATGTGGACAGCAGTTATTTTTTTATGAAAA  
TAGTATCACTGAACAAGCAGACAGTTAGTGAAGGATGCGTCAGGAAGCCTGCAGGCCACACAGCCATTTCTCTCGAAGAC  
15 TCCGGGTTTTCTGTGCATCTTTTGAACTCTAGCTCCAATTATAGCATGTACAGTGGATCAAGGTTCTTCTTCATTAA  
GGTTCAAGTTCTAGATTGAAATAAGTTTATGTAAACAGAAACAAAATTTCTGTACACACAACCTGCTCTGGGATTGGA  
GGAAAGTGTCTCGAGCTGGCGGCACACTGGTCAAGCCTCTGGGACAGGATACCTCTGGCCCATGGTCTATGGGGCGCTGG  
GCTTGGGCCCTGAGGGTACACAGTGCACCATGCCAGCTTCTGTGGATAGGATCTGGGTCTCGGATCATGCTGAGGACC  
ACAGCTGCCATGCTGGTAAAGGGCACCACGTGGCTCAGAGGGGGCGAGGTTCCAGCCCCAGCTTTCTTACCGTCTTCAG  
20 TTATTTTTCCCTAAGAGTCTGAGAAGTGGGGCCGCGCTGATGGCCTTCGTTCTGTCTTCAGCTGGCAGAGATTGCACAA  
GCTGATGGTAAACACTGAGTACTTATAATGAATGAGGAATTGCTGTAGCAGTTAACTGTAGAGAGCTCGTCTGTTGAAA  
GAAATTTAAGTTTTTCAATTAACCGCTTTGGAGAATGTTACTTTAATTTATGGCTGTGTAATTTGTTTGACATTCAGTCCC  
TCGTAGACAGATACTACGTAAAAAGTGTAAGTTAACCTTGCTGTGTATTTTCCCTATTTTAG

25 **Intron 10 (SEQ ID NO 14)**

GTGAGGCCCGTGCCGTGTGTCTGTGGGGACCTCCACAGCCTGTGGGCTTTGCAGTTGAGCCCCCGTGTCTGCCCTGG  
CACCGCAGCGTTGTCTCTGCCAAGTCCTCTCTCTCTGCCGCTGCTGGATCCGCAAGAGCAGAGGCGCTTGGCCGTGCACC  
CAGGCCTGGGGGCGCAGGGGCACCTTCGGGAGGGAGTGGGTACCGTGCAGGCCCTGGTCTGCAGAGACGCACCCAGGTT  
ACACACGTGGTGAGTGCAGGCGGTGACCTGGCTCCTGCTGCTCTTTGGAAAGTCAAGAGTGGCGGCTCCTGGGGCCCCAG  
30 TGAGACCCCCAGGAGCTGTGCACAGGGCCTGCAGGGCCGAGGCGGCAGCCTCCTCCCCAGGGTGCACCTGAGCCTGCGGA  
GAGCAGGAGCTGCTGAGTGAGCTGGCCACAGCGTTTCGCTGCGGTACGTTTCTGCGTGGGGTTGTTTGGGATCGGTGGG  
AGAATTTGGATTTGCTGAGTGTCTGTCTTGAACACGGAGATGGCTAGGAGTGGGTTTCAGAGTTGATTTTGTGAAT  
CAAATAAAATCAGGCACAGGGGACCTGGCCTCAGCACAGGGGATTGTCCAATGTGGTCCCCCTCAAGGGCGCCCCACAG  
AGCCGGTGGGCTTGTTTTAAAGTGCATTTTACGAGGGACGAGAAACCTTGAAAGCTGTAAAGGGAACCTCAGAAAATG  
35 TGGCCGCCAGGGGTGGTTTCAGGTGCTTTGCTGGGCTGTGTTTGTGAAAACCCATTTGGACCCGCCCTCCAAGTCCACCC  
TCCAGGTCCACCCTCCAGGGCCGCCCTGGGCTGGGGGTATGCCTGGCGTTCTTGTGCCGACCCCGAGCACAGCAGGC  
TGTGCACATTTAAATCCACTAAGATTCACTCGGGGGAGCCAGGTCCCAAGCAACTGAGGGCTCAGGAGTCTGAGGCT  
GCTGAGGGGACAGAGCAGACGGGAACGCTGCTTCTGTGTGGCAAGTTCTGAGGGTGTGGCCAGGGAGGTGGCTCAGA  
GTGTATGTTGGGGTCCCACCGGGGCGAGAACTCTGTCTCTGATGAGTCGGCAGCCATGTAACAGGAAGGGGTGGCCACAG  
40 GGAGCTGGGAATGCACCAGGGGAGCTGCGCAGCTGGCCGAGGTCCCAGGGCCAGGCCACAGGAAGGGCAGGGGGACGCC  
GGGGCCACAGCAGAGGCCGAGGAAGGGAAGGGGATGCCAGGCCAGAGCAGAGGCTACCGGGCAGGGGGGCTCCCTG  
AGCTGGGTGAGCGAGGCTCATGACTCGGCGAGGGAACCTCCTTGACGTGAAGCTGACGACTGGTGTGCCCCACTCACAG  
CCCAGCCAGGTCCCGCGCCTGAGCAGGAACCTCAGAACCTCCCTTTGTCTAAAGCACAGCAGATGCCTTCAGGGCATCT  
AGGAGAAAACAGGCAAAGTCGTTGAGAAACGCTCTAAAGAAGGTGGGATGGTGGCAATTTCTGTCCAGATTTAGTCT  
45 GCCCGGACCACAGATGAGTCTATAACGGGATGTGGTGTGGCATGGGGACACATGAGATGGACCATCACAGAGGCCAC  
TGGGGCTGCACCTCCCATCTGAGTCTGGCTGTCCCGGGTCCAGGCCAGGTTCTTGCACTGCTCACCTACCTGCTCCTGCC

GGGAGACAGGGAAAGCACCCCGAAGTCTGGAGCAGGGCTGGGTCCAGGCTCCTCAGAGCTCCTGCCAGGCCCAGCACCCCT  
GCTCCAAATCACCCTTCTCTGGGGTTTTCCAAAGCATTTAACAAGGGTGTACAGTTACCTCCTGGGTGACGGCCCCGCA  
TCCTGGGGCTGACATTGCCCTCTGCCTTAG

5 **Intron 11 (SEQ ID NO 15)**

GTGAGCGACCTGGCCGGAAGTGGAGCCTGTGCCCGGCTGGGGCAGGTGCTGCTGCAGGGCCGTTGCGTCCACCTCTGCT  
TCCGTGTGGGGCAGGCGACTGCCAATCCCAAAGGGTCAAGGGCCACAGGGTGCCCTCGTCCCATCTGGGGCTGAGCAGA  
AATGCATCTTTCTGTGGGAGTGAGGGTGCTCACAACGGGAGCAGTCTTTCTGTGCTATTTTGGTAAAAGGAAATGGTGCAC  
CAGACCTGGGTGCACTGAGGTGTCTTCAGAAAGCAGTCTGGATCCGAACCCAGACGCCCGGGCCCTGCTGGGCGTGAGT  
10 CTCTCAAACCCGAACACAGGGGCCCTGCTGGGCATGAGTCCCTCTGAACCCGAGACCTGGGGCCCTGCTGGGCGTGAGT  
CTCTCCGAACCCAGAGACTTCAGGGCCCTTTTGGGCGTGAGTCTCTCCGCTGTGAGCCCCACACTCCAAGGCTCATCCAC  
AGTCTACAGGATGCCATGAGTTCATGATCACGTGTGACCCATCAGGGGACAGGGCCATGGTGTGGGGGGGTCTCTACAA  
AATTCTGGGGTCTTGTTCCTCCAGAGCCCGAGAGCTCAAGGCCCGTCTCAGGCTCAGACACAAATGAATTGAAGATGGA  
CACAGATGCAGAAATCTGTGCTGTTTCTTTATGAATAAAAAGTATCAACATTCCAGGCAGGGCAAGGTGGCTCACACCT  
15 ATAATCCCAGCACTTTGGGAGGCCGAGGTGGGTGGATCATTGAGGCCAGGAGTTTGAAGCCAACCTAACCAACATAGTG  
AAATTCATTTCTACTTAAAAAATACAAAAATAGCCTGGCCTGGTGGCACACGCCTGTAGTCCCCGCTATGCGGGAGGC  
TGAGGCAGGAGAATCATTGAACCCAGGAGGCAGAGGTGTCAGTGAGCCGAGATCACACCACTGCACTCCAGCCTGGGCA  
ACAGAGTGAGACTTCATCTTAAAAAAGTATCAGCATTCCAAAACCATAGTGACAGGTGTTTTTTTATTC  
TGTCCTTCGATAATATTTACTGGTGTGTGTAGAGGCCGGAAGTGGGGGTGCCTTCCTCTGAAAGGCACACCTTCATGG  
20 GAAGAGAAATAAGTGGTGAATGGTTGTTAAACCAGAGGTTTAACTGGGGTCTGTGCTTCTGAGTTAACAGTCCAGATC  
TGGACTTTGCCTCTTTCCAGAATGCTCCCTGGGGTTTGTCTCATGGGGAGCAGCAGGTGTGGACACCCTCGTGATGGGG  
GAGCAGCAGGTGCAGACGCCCTCATGATGGGGGAGTGGCAGGTGCAGACACCCTGTGCATGGTGGCCAGCATGTCCCTG  
TTGCAGTCCCTCCCAAGGATGCCGGTCTCCTGTGCTCCCAAGTCCCTGCTTCCTCTCACAGCCTTACCTGGTC  
CTGGCCTCCACTGGCTTTGTCTGCATGATTTCCACATTTCTGGGCTCCAGCACCTCTTCGCTCTCCAGGCACCTCT  
25 GCAGTCTGGCCATACCAGTCAGCTGTGAAGTGTCCACTGCTTATTTTGTCTCCCATGAAATGTATTTTTTAGGACAGGC  
ACCCCTGGTTCCAGCCTCTGGCAGCAGCATCAGTGAATGTTATGAAGGACAAAGGACAGACAAACAAATCAGGAAAATGG  
GTTCTCTCTAAACACATTGCAAAGCCACAGAGGCTAGTGCAGGATGGGTGGGCATCAGGTGCATCAGATGTGGGTCCAATG  
CCAGAATATTCTGTGCTCCCAAAGGCCACTTGGTCAGAGTGTGTGCTTGCAGAGGTGGCTCTAAAAGCTCAGCAGTGGAG  
GCAGTGGTTCGCCATACTCAGGGTGAAGTCCATCCTCTGTGTCTGAAGTATACAGCAGAGGCTTGAAGGGCATCTGGGA  
30 GAAGAAAACAGGCAAAATGATTAAGAAAAGTAAAAAGGAAAAGTGGTAAGATGGGAATTTTCTTGTCCAGATTTTAGTC  
TCCCAAACCCAGCTCAGATGGTAGAATGTGGTCAGAACTGATGGACAGAACAAATAGAACAAAACGGAAGCCCTATCTCT  
CAGAAAACGTGTGTTAATGTGGTATGTGGCACAGCTGATGGAAAAGAGAGTGTGTGTGTAATTTTTTTCTGAGAAAAC  
GACTGGAAGCAAATAAGTTGTGTCTTTACAGCATATACCAGAGCAGATTCTAGGTAGAAGAGGAGACACATGCAACAAC  
ACCAGCAACAGAAATAAAACAAAAGACTCAAAGGGAAGGGAGGTGAACGTTCCCTGGTTTGGTGTGGGGAAGGACACAC  
35 AGGGAGGCGGATGAAACCAAGTGGAGCAACGGGCATTGCTTTCACTGCAGAGAACTCAGCTTGCTGAGCCACAGTGAAA  
ATGGCCATTCCCTGGAGCGTTTGTGCACGTGATTTATTTAAGGCGCCCTGTGAGGTCTGCACATTATCCTCTCACTTT  
GTTCTCTAACCACCTGAGAGGTAGAGGAGGAAAGGCTCCAGGGGAGCAGCCGCCCTTGGTCACCCAGCTGGCAAAGGGC  
ATGCATGATTGCAGCCTGGCCTCCTGCTCCGGGGCCCTTGTCTGCCCCAGGACCCACACAAGTCAGACCCATAGGCTC  
AGGGTGAGCCGGAGCCCAAGGTCGTGTTGGGGATGGCTGTGAAAGAAGAAATGGACGCTCTGATGCACACTTGGGAAGGTC  
40 CTACCAGCAGCGTCAAAGAAATGCATGTGAACTGACAGCGAGACCCATCCCTCAAAGAAACGCACGTGAAACTGATGGC  
GAGACCTGTCCCCATCCCTCATGCTGGCTCCTTTTCTGGGCTTGCCAAGAGCCAGCATCAGGTTGAGGCAAGCTGGAAAG  
ACTTTTCTGGAAGCAGCTTGTGTTGCATGGAAGTCTCACAATGTCTGTGTCTTCCAGTAATTCCACTTCTGAAGTGA  
CCAGACATTATCAGGGTCTTATTTACCATTTCAGTGTTCAGGCAGGGGACTTGCCACAGCAAGTCACGAACCTGCC  
CAAATACAGGGCTAAGGAGATATTATGCATCACAAAACCTTGCTCTGCCATTAAACATTTTTCAAAGAATTTTGAAGAAT  
45 GTTTAATGGCACAAAACGTTTATTTCAATGTAGCAGTGTCAAAGCTGGATGTAAAAGAACACACCCAGGAGCCTGCCG  
TGAATGTCTATGTGTGTTTCTTTGGACATGGACATACATGGGCAGTGAGTGGTGGTGAGGCCCCTGGAGGACATCGGTGG

001260 942356

5

## 10

15

## 25

35

40

45

CCGGGACCTTAGGCTTATTTATTTGTTTAAAAACATTCTGGGCCTGGCTTCCGTTGTTGCTAAATGGGGAAAAGACATCC  
CACCTCAGCAGAGTTACTGAGAGGCTGAAACCGGGGTGCTGGCTTGACTGGTGTGATCTCAGGTCATTCCAGAAGTGGCT  
CAGGAAGTCAGTGAGACCAGGTACATGGGGGGCTCAGGCAGTGGGTGAGATGAGGTACACGGGGGGCTCAGGCAGTGGGT  
GAGGCCAGGTACATGGGGGGCTCAGGCAGTGGGTGAGATGAGGTACACGGGGGGCTCAGGCAGAGGGTCAGACCAGGTAC  
5 ACGGGGGCTCTGATCACACGCACATATGAGCACATGTGCACATGTGCTGTTTCATGGTAGCCAGGTCTGTGCACACCTGC  
CCCAAAGTCCCAGGAAGCTGAGAGGCCAAAGATGGAGGCTGACAGGGCTGGCGCGGTGGCTCACACCTGTAGTCCCAGCA  
CTTTGGGAGGCCGAGGCGAGAGGATCCCTTGAGCCCAGGAGTTTAAAGACCAGCCTGAGCAACATAGTAGAACCCCATCTC  
TATGAAAAATAAAAACAAAAATTAGCTGAACATGGTGGTGTGCGCCTGTAGTTCCAATACTTGGGAGGCTGAAGTGGGAG  
GATCACTTGAGCCCAGGAGGTGGAAGCTGCAGTGAGCTGAGATTGCACCACTGTACTGCAGCCTGGGTGACAGAGTGAGA  
10 GCCCATCTCAACAACAACAAAGAAGACTGACAAATGCAGTTTCTTGAAAGAAACATTTAGTAGGAACTTAACCTACACA  
CAGAAGCCAAGTCGGTGTCTCGGTGTGAGTGAGATGATGGTCTCACACCATCACCCAGACCCAGGGTTTATG  
CACCACAGGGGCGGTGGCTCAGAAGGGATGCGCAGGACGTTGATATACGATGACATCAAGGTTGTCTGACGAAGGGCAG  
GATTCATGATAAGTACCTGCTGGTACACAAGGAACAATGGATAAACTGGAAACCTTAGAGGCCTTCCCGAACAGGGGCT  
AATCAGAAGCCAGCATGGGGGGCTGGCATCCAGGATGGAGCTGCTTCAGCCTCCACATGCGTGTTCATACAGATGGTGCA  
15 CAGAAACGCAGTGACCTGTGCACACACAGACACGCAGCTACTCGCACACACAAGCACACACAGACATGCATGCATGC  
ATCCGTGTGTGTGCACCTGTGCCCATGAGGAAACCCATGCATGTGCATTTCATGCACGCACACAGGCACCGGTGGGCCCAT  
GCCCACACCCACGAGCACCCTCTGATTAGGAGGCCTTTCCTCTGACGCTGTCCGCCATCCTCTCAG

**Intron 14 (WEQ ID NO 18)**

GTATGTGCAGGTGCCTGGCCTCAGTGGCAGCAGTGCCTGCCTGCTGGTGTAGTGTGTCAGGAGACTGAGTGAATCTGGG  
CTTAGGAAGTTCTTACCCCTTTTCGCATCAGGAAGTGGTTTAAACCAACCACTGTCAGGCTCGTCTGCCCCCCTCTCGT  
GGGGTGAGCAGAGCACCTGATGGAAGGGACAGGAGCTGTCTGGGAGCTGCCATCCTTCCACCTTGCTCTGCCTGGGGAA  
GCGCTGGGGGGCCTGGTCTCTCCTGTTTGCCCCATGGTGGGATTGGGGGGCCTGGCCTCTCCTGTTTGCCCTGTGGTGG  
GATTGGGCTGTCTCCCGTCCATGGCACTTAGGGCCCTTGTGCAAACCCAGGCCAAGGGCTTAGGAGGAGGCCAGGCCAG  
25 GCTACCCACCCCTCTCAGGAGCAGAGGCCGCTATCACCACACAGAGCCCCGCGCGTCTCTGCTTCCAGTCACCG  
TCCTCTGCCCCCTGGACACTTTGTCCAGCATCAGGGAGGTTTCTGATCCGTCTGAAATTCAAGCCATGTCGAACCTGCGGT  
CCTGAGCTTAACAGCTTCTACTTTCTGTTCTTTCTGTGTTGTGGAAATTTACCTGGAGAAGCCGAAGAAAACATTTCTG  
TCGTGACTCCTGCGGTGCTTGGGTGCGGACAGCCAGAGATGGAGCCACCCCGCAGACCGTGGGGTGTGGGCAGCTTTCCG  
GTGTCTCCTGGGAGGGGAGCTGGGCTGGGCCTGTGACTCCTCAGCCTCTGTTTTCCCCCAG

**Intron 15 (SEQ ID NO 19)**

GCAAGTGTGGGTGGAGGCCAGTGCGGGCCCCACCTGCCCAGGGGTATCCTTGAACGCCCTGTGTGGGGCAGCAGCCTC  
AGATGCTGCTGAAGTGCAGACGCCCCCGGGCCTGACCCTGGGGGCTGGAGCCACGCTGGCAGCCCTATGTGATTAAACG  
CTGGTGTCCCCAGGCCACGGAGCCTGGCAGGGTCCCCAACTTCTGAACCCCTGCTTCCCATCTCAGGGGCGATGGCTCC  
35 CCACGCTTGGGAGCCTTCTGACCCCTGACCTGTGTCTCTACAGCCTCTTCCCTGGCTGCTGCCCTGAGCTCCTGGGGT  
CCTGAGCAAGTTCTCTCCCCGCCCCGCGCTCCAGCGTCACTGGGTGCTGTCTGCTCGCCCCGGTGGAGGGGTGTCTG  
TCCCTTCACTGAGGTTCCCACAGCCAGGGCCACGAGGTGCAGGCCCTGCCTGCCCGGCCACCCACACGTCCTAGGAGGG  
TTGGAGGATGCCACCTCTGGCCTCTTCTGGAACGGAGTCTGATTTTGGCCCCGAG

**3'-untranscribed region (SEQ ID NO 20)**

ATCTCATGTTTGAATCCTAATGTGCACTGCATAGACACCACTGTATGCAATTACAGAAGCCTGTGAGTGAACGGGGTGGT  
GGTCAGTGCGGGGCCATGGCCTGGCTGTGCATTTACGGAAGTCTATGAGTGAATGGGGTTGTGGTCAGTGGGGGCCATG  
GCCTGGCTGGGCCTGGGAGGTTTCTGATGCTGTGAGGCAGGAGGGGAAGGAGGGTAGGGGATAGACAGTGGGAGCCCCCA  
CCCTGGAAGACATAACAGTAAGTCCAGGCCCGAAGGGCAGCAGGGATGCTGGGGGCCAGCTTGGGCGGCGGGGATGATG  
45 GAGGGCCTGGCCAGGGTGGCAGGGATGATGGGGGCCCGAGCTGGGGTGGCAGGGGTGATGGGGGGGGCTGGTCTGGGTGG

CGGGGAAGATGGGGAAGCCTGGCTGGGCCCCCTCCTCCCCTGCCTCCACCTGCAGCCGTGGATCCGGATGTGCTTCCCT  
GGTGCACATCCTCTGGGCCATCAGCTTTCATGGAGGTGGGGGCGAGGGGCATGACACCATCCTGTATAAAATCCAGGATT  
CCTCCTCCTGAACGCCCCAACTCAGGTTGAAAGTCACATTCCGCCCTCTGGCCATTCTCTTAAGAGTAGACCAGGATTCTG  
ATCTCTGAAGGGTGGGTAGGGTGGGGCAGTGGAGGGTGTGGACACAGGAGGCTTCAGGGTGGGGCTGGTGATGCTCTCTC  
5 ATCCTCTTATCATCTCCCAGTCTCATCTCTCATCTCTTATCATCTCCCAGTCTCATCTGTCTTCTCTTATCTCCCAGT  
CTCATCTGTATCCTCTTACCATCTCCCAGTCTCATCTCTTATCCTCTTATCTCCTAGTCTCATCCAGACTTACCTCCCA  
GGGCGGGTGCCAGGCTCGCAGTGGAGCTGGACATACGTCTTCTCAGGCAGAAGGAACTGGAAGGATTGCAGAGAACAG  
GAGGGGCGGCTCAGAGGGACGCAGTCTTGGGGTGAAGAAACAGCCCCCTCCTCAGAAGTTGGCTTGGGCCACACGAAACCG  
AGGGCCCTGCGTGAGTGGCTCCAGAGCCTTCCAGCAGGTCCCTGGTGGGGCCTTATGGTATGGCCGGGTCTACTGAGTG  
10 CACCTTGGACAGGGCTTCTGGTTTGTAGTGCAGCCCGACGTGCCTGGTGTGCGGGTGGGGGCTTATGGCCACTGGATATG  
GCGTCATTTATTGCTGCTGCTTCAAGAAATGTCTGAGTGACCGAGCCTAATGTGTATGGTGGGCCCCAAGTCCACAGACTG  
TGTCGTAATGCACTCTGGTGCCTGGAGCCCCGTATAGGAGCTGTGAGGAAGGAGGGGCTCTTGGCAGCCGGCCTGGGG  
GCGCCTTTGCCCTGCAAACCTGGAAGGGAGCGCCCCGGGCGCCGTGGGCGGACGACCTCAAGTGAGAGTTGGACAGAAC  
AGGGCGGGGACTTCCCAGGAGCAGAGGCCGCTGCTCAGGCACACCTGGGTTTGAATCACAGACCAACaGGTCAGGCCATT  
15 GTTCAGCTATCCATCTTCTACAAAGCTCCAGATTCTGTCTTCTCCGGGTGTTTTTTTGTGAAATTTTACTCAGGATTACT  
TATATTTTTTGTAAAGTATTAGACCCTTAAAAAAGGTATTTGCTTTGATATGGCTTAACCTACTAAGCACCTACTTTAT  
TTGTCTGTTTTTATTATTATTATTATTATTATTAGAGATGGTGTCTACTCTGTACCCAGGTTGTTAGTGCAGTGGCAC  
AGTCATGGCTCGCTGTAGCCGCAAACCCCAGGCTCAAGTGATCCTCCGGCCTCAGCTTCCCAGAGTGTGGGATTACAG  
GTGTGAGCCACTGCCCTTGCTGGCACTTTTAAAAACCCTATGTAAGGTGAGTCCAGTGGCTTCCACACCTGTCATCC  
20 CAGTAGTTTGGGAAGCCGAGGCAGAAGGATTGTCTGAGGCCAGGAGTTTGTAGACCAGCATGGGTAACATAGGGAGACCCC  
ATCTCTACAAAAAATGCAAAAAGTTATCCGGGCGTGGGGTCCAGCATCTGTAGTCCCAGCTGCTCGGGAGGCTGAGTGGG  
AGGATCGCTTGAGCCCGGAGGTGATGGCTGCAGTGAGCTGTGATTGTACCATCGCACTCCAGCCTGGGCAACAGAGTGA  
GACCTGTCTCAAAAAAAAAAAAAAAAAAAGAGGAGAAGGAGAAGAGAAGGAAGGAAGGAAGAGAAGAGAAG  
GAAGAAGGAAGAAAGGAAGGAGAAGGAGGCTGCTAGGTGCTAGGTAGACTGTCAAATCTCAGAGCAAAATGAAAATAACA  
25 AAGTTTTAAAGGGAAGAAAAACCCAGCTCTTTGGACTTCCTTAGGCCTGAACCTCATCTCAAGCAGCTTCCCTCCACA  
GACAAGCGTGATGGAGCGAGTGAGTTCAAAGCAGAAAGGGAGGAGAAGCAGGCAAGGGTGGAGGCTGTGGGTGACACCA  
GCCAGGACCCCTGAAAGGGAGTGGTTGTTTTCTGCCTCAGCCCCACGCTCCTGCCGGTCTGCACCTGCTGTAACCGTC  
GATGTTGGTGCCAGGTGCCACCTGGGAAGGATGCTGTGCAGGGGGCTTGCCAACTTTGGTGGGTTTCAGAAGCCCCAG  
GCACTTGTGGCAGGCACAATTACAGCCCCCTCCCCAAAGATGCCACGTCCTTCTCCTGGAACCTGTGAATGTGTACCCG  
30 CAAGGCAGAGGCTGGTGAAGGCTGCAGGTGGAATCACGGCTGCCAGTCAGCCGATCTTAAGGTCATCCTGGATTATCTGG  
TGGGCCTGATATGGCCACAAGGGTCCCTAGAAGTGAGAGAGGGAGGCAGGGGAGAGTCAGAGAGGGGACGTGAGAAGGAC  
CACTGGCCACTGCTGGCTTTGAGATGGAGGAGGGGGTCCCCAGCCAAGGAATGGGGGCAGCCGCTCCATGCTGGAAAAGC  
AAGCAATCCTCCCCGGTCTTGAGGGCACACGGCCCTGCCACGCTCGATTTCAGGCCAGTGGGACCTGTTTCAGCTTTC  
CGGCCTCCAGAGCTGTAAGATGATGCGTTTGTGTTAGCCACTAAGCTGCAGTGATTCGTACAGCAGCAATGGAATAG  
35 CAGTACAGGGAAATGAATACAGGGACAGTTCTCAGAGTGACTCTCAGCCCACCCCTGGG

001260" 912360



Characterization of the exons showed, interestingly, that the functionally important hTC protein domains which are described in our Patent Application PCT/EP/98/03469 are arranged on separate exons. The telomerase-characteristic T motif is located on exon 3. The RT (reverse transcriptase) motifs 1-7, which are important for the catalytic function of the telomerase, are located on the following exons: RT motifs 1 and 2 on exon 4, RT motif 4 on exon 9, RT motif 5 on exon 10, and RT motifs 6 and 7 on exon 11. RT motif 3 is shared by exons 5 and 6 (see Fig. 8).

Elucidation of the exon-intron structure of the hTC gene also shows that the four deletions or insertion variants of the hTC cDNA which were described in our Patent Application PCT/EP/98/03469, as well as three additional hTC insertion variants which are described in the literature (Kilian et al., 1997), in all probability represent alternative splicing products. As shown in Fig. 8, the splicing variants can be divided into two groups: deletion variants and insertion variants.

The hTC variants in the deletion group lack specific sequence segments. The 36 bp in-frame deletion in variant DEL1 in all probability results from using an alternative 3' splice acceptor sequence in exon 6, resulting in a part of RT motif 3 being lost. In variant DEL2, the normal 5' splice donor and 3' splice acceptor sequences of introns 6, 7 and 8 are not used. Instead exon 6 is fused directly to exon 9, resulting in a displacement arising in the open reading frame and a stop codon appearing in exon 10. Variant Del3 is a combination of variants 1 and 2.

The insertion variant group is characterized by the insertion of intron sequences which lead to premature cessation of translation. Instead of the 5' splice donor sequence of intron 5, which is normally used, use is made, in variant INS1, of an alternative, 3'-located splice site, resulting in the insertion of the first 38 bp from intron 4 between exon 4 and exon 5. The insertion, in variant INS2, of a region of the intron 11 sequence likewise results from using an alternative 5' splice donor sequence in intron 11. Since this variant was only described inadequately in the

literature (Kilian et al., 1997), it is not possible to determine the precise alternative 5' splice donor sequence in this variant. The insertion of intron 14 sequences between exon 14 and exon 15 in variant INS3 comes from using an alternative 3' splice acceptor sequence, resulting in the 3' part of intron 14 not being spliced.

5

The hTC variant INS4 (variante 4), which is described in our Patent Application PCT/EP/98/03469, is characterized by exon 15, and the 5' part region of exon 16, being replaced by the first 600 bp of intron 14. This variant can be attributed to the use of an alternative internal 5' splice donor sequence in intron 14 and an alternative 3' splice acceptor sequence in exon 16, resulting in an altered C terminus.

10

The *in vivo* generation of hTC protein variants which are probably non-functional and which could interfere with the function of the complete hTC protein constitutes a possible mechanism, in addition to transcription regulation, for controlling hTC protein function. The function of the hTC splicing variants is not yet known. Although most of these variants presumably encode proteins without reverse transcriptase activity, they could nevertheless play a crucial role as transdominant-negative telomerase regulators by, for example, competing for interaction with important binding partners.

15

20

The search for possible transcription factor binding sites was carried out using the „find pattern“ algorithm from the Genetics Computer Group (Madison, USA) GCG Sequence Analysis program package. This resulted in the identification of a variety of potential binding sites for transcription factors in the nucleotide sequence of intron 2, which binding sites are listed in Tab. 2. In addition, an Sp1 binding site was found in intron 1 (pos. 43), and a c-Myc binding site was found in the 5'-untranslated region (cDNA position 29-34, cf. Fig. 6).

25

007260-3423360

**Example 6**

In order to ascertain the start point(s) of hTC transcription in HL 60 cells, the 5' end of the hTC mRNA was determined by means of primer extension analysis.

5

2 µg of polyA<sup>+</sup> RNA from HL-60 cells were denaturated at 65°C for 10 min. 1 µl of RNasin (30-40 U/ml) and 0.3-1 pmol of radioactively labelled primer (5'GTTAAGTTGTAGCTTACACTGGTTCTC 3'; 2.5-8x10<sup>5</sup> cpm) were added for primer annealing, and the whole was incubated, at 37°C for 30 min, in a total volume of 20 µl. After the addition of 10 µl of 5xreverse transcriptase buffer (from Gibco-BRL), 2 µl of 10 mM dNTPs, 2 µl RNasin (see above), 5 µl of 0.1 M DTT (from Gibco-BRL) 2 µl of ThermoScript RT (15 U/µl; from Gibco-BRL) and 9 µl of DEPC-treated water, primer extension took place, at 58°C for 1 h, in a total volume [lacuna]. The reaction was stopped by adding 4 µl of 0.5 M EDTA, pH 8.0, and the RNA was degraded, at 37°C for 30 min, after having added 1 µl of RNaseA (10 mg/ml). 2.5 µg of sheared calf thymus DNA and 100 µl of TE were then added, and the mixture was extracted once with 150 µl of phenol/chloroform (1:1). The DNA was precipitated, at -70°C for 45 min, after adding 15 µl of 3 M Na acetate and 450 µl of ethanol, and then centrifuged at 14,000 rpm for 15 min. The precipitate was washed once with 70% ethanol, dried in air and dissolved in 8 µl of sequencing stop solution. After 5 min of denaturation at 80°C, the samples were loaded onto a 6% polyacrylamide gel and fractionated electrophoretically (Ausubel et al., 1987) (Fig. 5).

25 In this connection, a main transcription start site was identified which is located 1767 bp 5' of the ATG start codon of the hTC cDNA sequence (nucleotide position 3346 in Fig. 4). In addition to this, the nucleotide sequence around this main transcription start (TTA<sub>+1</sub>TTGT) represents an initiator element (Inr), which, in 6 out of 7 nucleotides, matches the consensus motif (PyPyA<sub>+1</sub>Na/tPyPy) (Smale, 1997) of an initiator element.

30

It was not possible to identify any unambiguous TATA box in the immediate vicinity of the experimentally identified main transcription start, which means that the hTC promoter has probably to be classified in the family of TATA-less promoters (Smale, 1997). However, a potential TATA box from nucleotide position 1306 to nucleotide position 1311 (Fig. 4) was found by means of bioinformatics analysis. The subsidiary transcription starts which were additionally observed around the main transcription start have also been described in the case of other TATA-less promoters (Geng and Johnson, 1993), for example in the strongly regulated promoters of some cell cycle genes (Wick *et al.*, 1995).

#### Example 7

In addition to the start point of the hTC transcript which was described in Example 6 and identified in HL60 cells, a further transcription start region was also identified in HL60 cells. With the aid of RT-PCR analyses, the region of the hTC gene transcription start in HL60 cells was localized to bp -60 to bp -105.

The cDNA for this was synthesized using a First Strand cDNA Synthesis kit (Clontech), in accordance with the manufacturer's instructions, and employing 0.4 µg of HL60 cell polyA RNA (Clontech) and the gene-specific primer GSP13 (5'-CCTCCAAAGAGGTGGCTTCTTCGGC-3', cDNA position 920-897). In a final volume of 50 µl, 10 pmol dNTP mix were added to 1 µl of cDNA, and a PCR reaction was carried out in 1xPCR reaction buffer F (PCR-Optimizer kit from InVitrogen) and using one unit of platinum Taq DNA polymerase (from Gibco/BRL). 10 pmol of each of the 5' and 3' primers defined below were added as primers. The PCR was carried out in 3 steps. A two-minute denaturation at 94°C was followed by 36 PCR cycles in which the DNA was first of all denatured at 94°C for 45 sec and, after that, the primers were annealed, and the DNA chain was extended at 68°C for 5 min. The cycles were concluded by a chain extension at 68°C for 10 min. In all, six different 5' PCR primers (primer HTRT5B: 5'-CGCAGCCACTACCGCGAGGTGC-3', cDNA position 105 to 126; primer C5S:

5'-CTGCGTCCTGCTGCGCACGTGGGAAGC-3', 5'-flanking region -49 to -23; primer PRO-TEST1: 5'-CTCGCGGCGCGAGTTTCAGGCAG-3', 5'-flanking region -74 to -52; primer PRO-TEST2: 5'-CCAGCCCCTCCCCTTCCTTTCC-3', 5'-flanking region -112 to -91; primer PRO-TEST4: 5'-CCAGCTCCGCCTCCTCCGCGC-3', 5'-flanking region -191 to -171; primer RP-3A: 5'-CTAGGCCGATTCGACCTCTCTCC-3', 5'-flanking region -427 to -405) were combined with the 3' PCR primer C5Rback (5'-GTCCCAGGGCACGCACACCAG-3', cDNA position 245 to 225). Genomic DNA was also employed for the PCR, as a control, in addition to the Oligo dT- and GSP13-primed cDNAs. As Fig. 9 shows, a PCR product was only obtained with the primer combinations HTRT5B-C5Rback, C5S-C5Rback and PRO-TEST1-C5Rback, indicating that the start point for hTC transcription lies in the region between bp-60 and bp-105.

### Example 8

Several extremely GC-rich regions, so-called CpG Islands, are located in the isolated 5'-flanking region, of about 11.2 kb in size, of the hTC gene. One CpG Island, having a GC content of > 70%, extends from bp - 1214 into intron 2. Two further GC-rich regions having a GC content of > 60% extend from bp -3872 to bp -3113 and from bp -5363 to bp -3941, respectively. The positions of the CpG Islands are shown graphically in Fig. 11.

The search for possible transcription factor binding sites was carried out using the "Find Pattern" algorithm from the Genetics Computer Group (Madison, USA) GCG Sequence Analysis program package. This resulted in the identification of a variety of potential binding sites in the region up to -900 bp upstream of the translation start codon ATG: five Sp1 binding sites, one c-Myc binding site, and one CCAC box (Fig. 10). In addition, a CCAAT box and a second c-Myc binding site were found at positions -1788 and -3995, respectively, of the 5'-flanking region.

**Example 9**

In order to analyse the activity of the hTC promoter, PCR amplification was used to generate four hTC promoter sequence segments of differing length, which segments were cloned into the Promega vector pGL2 5' in front of the luciferase reporter gene. The 8.5 kb SacI fragment which was subcloned from phage clone P12 was selected as the DNA source for the PCR amplification. In a final volume of 50 µl, 10 pmol of dNTP mix were added to 35 ng of this DNA, and a PCR reaction was carried out in 1xPCR reaction buffer (PCR-Optimizer kit from InVitrogen) and using one unit of platinum Taq DNA polymerase (from Gibco/BRL). In each case 20 pmol of the 5' and 3' primers which are defined below were added as primers. The PCR was carried out in three steps. A two-minute denaturation at 94°C was followed by 30 PCR cycles in which the DNA was first of all denaturated at 94°C for 45 sec, after which the primers were annealed, and the DNA chain was extended, at 68°C for 5 min. The cycles were concluded by a chain extension at 68°C for 10 min. The selected 3' PCR primer was in each case the primer PK-3A (5'-GCAAGCTTGACGCAGCGCTGCCTGAACTCG-3', position -43 to -65), which primer recognizes a sequence region 42 bp upstream of the ATG START codon. A promoter fragment of 4051 bp in size (NPK8) was amplified by combining the PK-3A primers with the 5' PCR primer PK-5B (5'-CCAGATCTCTGGAACACAGAGTGGCAGTTTCC-3', position -4093 to -4070). Combining the pair of primers PK-3A and PK-5C (5'-CCAGATCTGCATGAAGTGTGTGGGGATTTGCAG-3', position -3120 to -3096) led to the amplification of a promoter fragment of 3078 bp in size (NPK15). Use of the primer combination PK-3A and PK-5D (5'-GGAGATCTGATCTTGGCTTACTGCAGCCTCTG-3', position -2110 to -2087) amplified a promoter fragment of 2068 bp in size (NPK22). Finally, using the primer combination PK-3A and PK-5E (5'-GGAGATCTGTCTGGATTCCTGGGAAGTCCTCA-3', position -1125 to -1102) led to the amplification of a promoter fragment of 1083 bp in size (NPK27).

The PK-3A primer contains a HindIII recognition sequence. The different 5' primers contain a BglII recognition sequence.

5 The resulting PCR products were purified using the Qiagen QIA quick spin PCR purification kit, in accordance with the manufacturer's instructions, and then digested with the restriction enzymes BglII and HindIII. The pGL2 promoter vector was digested with the same restriction enzymes, and the SV40 promoter contained in this vector was released and removed. The PCR promoter fragments ligated into the vector, which was then transformed into competent DH5 $\alpha$  bacteria (from  
10 Gibco/BRL). DNA for the promoter activity analyses, which are described below, was isolated from transformed bacterial clones using the Qiagen plasmid kit.

#### **Example 10**

15 The activity of the hTC promoter was analysed in transient transfections in eukaryotic cells.

All the work with eukaryotic cells was carried out at a sterile workstation. CHO-K1 and HEK 293 cells were obtained from the American Type Culture collection.

20 CHO-K1 cells were kept in DMEM Nut Mix F-12 cell culture medium (from Gibco-BRL, order number: 21331-020) containing 0.15% streptomycin/penicillin, 2 mM glutamine and 10% FCS (from Gibco-BRL).

25 HEK 293 cells were cultured in DMOD cell culture medium (from Gibco-BRL, order number: 41965-039) containing 0.15% streptomycin/penicillin, 2 mM glutamine and 10% FCS (from Gibco-BRL).

30 CHO-K1 and HEK 293 cells were cultured at 37°C in a water-saturated atmosphere while being gassed with 5% CO<sub>2</sub>. When the cell lawn was confluent, the medium was sucked off, after which the cells were washed with PBS (100 mM KH<sub>2</sub>PO<sub>4</sub> pH

7.2; 150 mM NaCl) and released by adding a trypsin-EDTA solution (from Gibco-BRL). The trypsin was inactivated by adding medium and the cell count was determined using a Neubauer counting chamber in order to plate out the cells at the desired density.

5

For the transfection, in each case  $2 \times 10^5$  HEK 293 cells were plated out, per well, in a 24-well cell culture plate. The HEK 293 medium was removed after 3 hours. For the transfection, up to 2.5  $\mu$ g of plasmid DNA, 1  $\mu$ g of a CMV  $\beta$ -Gal plasmid construct (from Stratagene, order number: 200388), 200  $\mu$ l of serum-free medium and 10  $\mu$ l of transfection reagent (DOTAP from Boehringer Mannheim) were incubated at room temperature for 15 minutes and then dropped uniformly onto the HEK 293 cells. 1.5 ml of medium were added after 3 hours. The medium was changed after 20 hours. After a further 24 hours, the cells were harvested for determining the luciferase activity and the  $\beta$ -Gal activity. For this, the cells were lysed, at room temperature for 15 minutes, in the cell culture lysis reagent (25 mM Tris [pH 7.8] containing  $H_3PO_4$ ; 2 mM CDTA; 2 mM DTT; 10% glycerol; 1% Triton X-100). Twenty  $\mu$ l of this cell lysate were mixed with 100  $\mu$ l of luciferase assay buffer (20 mM Tricin; 1.07 mM  $(MgCO_3)_4$   $Mg(OH)_2 \cdot 5H_2O$ ; 2.67 mM  $MgSO_4$ ; 0.1 mM EDTA; 33.3 mM DTT; 270  $\mu$ M coenzyme A; 470  $\mu$ M luciferin, 530  $\mu$ M ATP), and the light generated by the luciferase was measured.

In order to measure the  $\beta$ -galactosidase activity, equal quantities of cell lysate and  $\beta$ -galactosidase assay buffer (100 mM sodium phosphate buffer, pH 7.3; 1 mM  $MgCl_2$ ; 50 mM  $\beta$ -mercaptoethanol; 0.665 mg of ONPG/ml) were incubated at 37°C for at least 30 minutes or until a slight yellow coloration appeared. The reaction was stopped by adding 100  $\mu$ l of 1 M  $Na_2CO_3$ , and the absorption was determined at 420 nm.

In order to analyse the hTC promoter, four hTC promoter sequence segments of differing length were cloned 5' in front of the luciferase reporter gene (cf. Example 9).



5 The construct NPK 27 exhibits a luciferase activity which is 40 times higher than the basal activity of the promoterless luciferase control construct (pGL2-basic) and from 2 to 3 times higher than that of the SV40 promoter control construct (pGL2PRO). Interestingly, a luciferase activity which was from 2 to 3 times lower than that obtained with the NPK 27 construct was observed in cells which were transfected with longer hTC promoter constructs (NPK8, NPK15, NPK22). Similar results were  
10 also observed in CHO cells (data not shown).

## References

- 5      **Allsopp, R. C., Vazire, H., Pattersson, C., Goldstein, S., Younglai, E.V., Futcher, A.B., Greider, C.W. und Harley, C.B. (1992).** Telomere length predicts replicative capacity of human fibroblasts. Proc. Natl. Acad. Sci. **89**, 10114-10118.
- 10      **Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (1987).** Current protocols in molecular biology. Greene Publishing Associates and Wiley-Intersciences, New York.
- 15      **Blasco, M. A., Rizen, M., Greider, C. W. und Hanahan, D. (1996).** Differential regulation of telomerase activity and telomerase RNA during multistage tumorigenesis. Nature Genetics **12**, 200-204.
- 20      **Broccoli, D., Young, J. W. und deLange, T. (1995).** Telomerase activity in normal and malignant hematopoietic cells. Proc. Natl. Acad. Sci. **92**, 9082-9086.
- 25      **Counter, C. M., Avilion, A. A., LeFeuvre, C. E., Stewart, N. G. Greider, C.W. Harley, C. B. und Bacchetti S. (1992).** Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. EMBO J. **11**, 1921-1929.
- 30      **Feng, J., Funk, W. D., Wang, S.-S., Weinrich, S. L., Avilion, A.A., Chiu, C.-P., Adams, R.R., Chang, E., Allsopp, R.C., Yu, J., Le, S., West, M.D., Harley, C.B., Andrews, W.H., Greider, C.W. und Villeponteau, B. (1995).** The RNA component of human telomerase. Science **269**, 1236-1241.
- 35      **Geng, Y., and Johnson, L.F. (1993).** Lack of an initiator element is responsible for multiple transcriptional initiation sites of the TATA less mouse thymidine synthasse promoter. Mol. Cell. Biol **14**:4894.
- Goldstein, S. (1990).** Replicative senescence: The human fibroblast comes of age. Science **249**, 1129-1133.
- Harley, C.B., Futcher, A.B., Greider, C.W., 1990.** Telomeres shorten during ageing of human fibroblasts. Nature **345**, 458-460.

- Hastie, N. D., Dempster, M., Dunlop, M. G., Thompson, A. M., Green, D.K. und Allshire, R.C. (1990). Telomere reduction in human colorectal carcinoma and with ageing. *Nature* **346**, 866-868.
- 5     Hiyama, K., Hirai, Y., Kyoizumi, S., Akiyama, M., Hiyama, E., Piatyszek, M.A., Shay, J.W., Ishioka, S. und Yamakido, M. (1995). Activation of telomerase in human lymphocytes and hematopoietic progenitor cells. *J. Immunol.* **155**, 3711-3715.
- 10     Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C. B., West, M.D., Ho, P.L.C., Coviello, G.M., Wright, W.E., Weinrich, S.L. und Shay, J.W. (1994). Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011-2015.
- 15     Latchman, D.S. (1991). Eukaryotic transcription factors. Academic Press Limited, London.
- 20     Lingner, J., Hughes, T.R., Shevchenko, A., Mann, M., Lundblad, V. und Cech T.R. (1997). Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science* **276**: 561-567.
- 25     Lundblad, V. und Szostak, J. W. (1989). A mutant with a defect in telomere elongation leads to senescence in yeast. *Cell* **57**, 633-643.
- 30     McClintock, B. (1941). The stability of broken ends of chromosomes in *Zea mays*. *Genetics* **26**, 234-282.
- 35     Meyne, J., Ratliff, R. L. und Moyzis, R. K. (1989). Conservation of the human telomere sequence (TTAGGG)<sub>n</sub> among vertebrates. *Proc. Natl. Acad. Sci.* **86**, 7049-7053.
- Olovnikov, A. M. (1973). A theory of marginotomy. *J. Theor. Biol.* **41**, 181-190.
- Sandell, L. L. und Zakian, V. A. (1993). Loss of a yeast telomere: Arrest, recovery and chromosome loss. *Cell* **75**, 729-739.
- Shapiro, M.B., Senapathy, P., 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucl. Acids Res.* **15**, 7155-7174.
- Smale, S.T. and Baltimore, D. (1989). The „initiator“ as a transcription control element. *Cell* **57**:103-113.

**Smale, S.T. (1997).** Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. *Biochimica et Biophysica Acta* 1351, 73-88.

5      **Shay, J. W. (1997).** Telomerae and Cancer. Ciba Foundation Meeting: Telomeres and Telomerase. London.

10      **Vaziri, H., Dragowska, W., Allsopp, R. C., Thomas, T. E., Harley, C.B. und Landsdorp, P.M. (1994).** Evidence for a mitotic clock in human hematopoietic stem cells: Loss of telomeric DNA with age. *Proc. Natl. Acad. Sci.* **91**, 9857-9860.

15      **Wick, M., Härönen, R., Mumberg, D., Bürger, C., Olsen, B.R., Budarf, M.L., Apte, S. S. and Müller, R. (1995).** Structure of the human TIMP-3 gene and its cell-cycle-regulated promoter. *Biochemical Journal* 311, 549-554.

**Zakian, V. A. (1995).** Telomeres: Beginning to understand the end. *Science* **270**, 1601-1607.

001260:91223560

## SEQUENCE LISTING

<110> Bayer AG

5 <120> Regulatory DNA sequences from the 5i region of the gene  
for the human catalytic telomerase subunit, and  
their diagnostic and therapeutic use

10 <130> LeA32805-Foreign Countries

<140>  
<141>

15 <160> 20

<170> PatentIn Vers. 2.0

<210> 1  
<211> 5126  
<212> DNA  
20 <213> Homo sapiens

<400> 1

25	gagctctgaa	ccgtggaaac	gaacatgacc	cttgccctgcc	tgcttccctg	ggtgggtcaa	60
	gggtaaatgaa	gtgggtgtgca	ggaaatggcc	atgtaaatta	cacgactctg	ctgatgggga	120
	ccgttcccttc	catcattatt	catcttcacc	cccaaggact	gaatgattcc	agcaacttct	180
	tcgggtgtga	caagccatga	caaaactcag	tacaaacacc	actcttttac	taggcccaca	240
	gagcacgggc	cacacccctg	atatattaag	agtccaggag	agatgaggct	gctttcagcc	300
	accaggctgg	ggtgacaaca	gcggctgaac	agtctgttcc	tctagactag	tagaccctgg	360
30	caggcactcc	cccaaattct	agggcctggt	tgctgcttcc	cgagggcgcc	atctgccctg	420
	gagactcagc	ctggggtgcc	acactgaggg	cagccctgtc	tccacaccct	ccgcctccag	480
	gcctcagctt	ctccagcagc	ttcctaaacc	ctgggtgggc	cggtgtccag	cgctactgtc	540
	tcacctgtcc	cactgtgtct	tgtctcagcg	acgtagctcg	cacggttcct	cctcacatgg	600
	ggtgtctgtc	tccttcccca	acactcacat	gcgttgaagg	gaggagattc	tgcgccctcc	660
35	agactggctc	ctctgagcct	gaacctggct	cgtagggccc	gatgcagggt	cctggcgctc	720
	ggctgcacgc	tgacctccat	ttccaggcgc	tccccgtctc	ctgtcatctg	ccggggcctg	780
	ccggtgtggt	cttctgtttc	tgtgtctcct	tccacgtcca	gctgcgtgtg	tctctgcccg	840
	ctagggctct	gggggtttta	taggcatagg	acgggggcgt	ggtgggccag	ggcgctcttg	900
40	ggaaatgcaa	catttgggtg	tgaaagttag	agtgcctgtc	ctcacctagg	tccacgggca	960
	caggcctggg	gatggagccc	ccgccaggga	cccgcccttc	tctgcccagc	acttttctgc	1020
	ccccctccct	ctggaacaca	gagtggcagt	ttccacaagc	actaagcatc	ctcttcccaa	1080
	aagaccagcg	attggcaccc	ctggacattt	gccccacagc	cctgggaatt	cactgtgagg	1140
	cgcacatcat	gtacacactc	ccgtccacga	ccgacccccg	ctgttttatt	ttaatatgta	1200
45	caaagcaggg	aaatccctgc	taaaatgtcc	tttaacaaac	tggttaaaca	aacgggtcca	1260
	tcgcgcagg	ggacagtctc	tcacagtga	gaggaacatg	ccgtttataa	agcctgcagg	1320
	catctcaagg	gaattacgct	gagtcaaaac	tgccacctcc	atgggatacg	tacgcaacat	1380
	gctcaaaaag	aaagaatttc	accccatggc	aggggagtg	ttaggggggt	taaggacggg	1440
	gggggcggca	gctgggggct	actgcacgca	ccctttacta	aagccagttt	cctggttctg	1500
50	atgggtattg	ctcagttatg	ggagactaac	cataggggag	tggggatggg	ggaacccgga	1560
	ggctgtgcca	tctttgccat	gcccagtggt	cctgggcagg	ataatgtctc	agagatggcc	1620
	acgtctctgt	tcccccaaac	ctgtggacag	aacccgcctc	gccccagggc	ctttgcagg	1680
	gtgatctccg	tgaggaccct	gaggtctggg	atccttcggg	actacctgca	ggcccgaaaa	1740
	gtaatccagg	ggttctggga	agagggcggc	aggagggtca	gagggggggc	gcctcaggac	1800
55	gatggaggca	gtcagtctga	ggctgaaaag	ggagggaggg	cctcgagccc	aggcctgcga	1860
	gcgcctccag	aagctggaaa	aagcggggaa	gggacctcc	acggagcctg	cagcaggaag	1920
	gcacggctgg	cccttagccc	accaggggcc	atcgtggacc	tccggcctcc	gtgccatagg	1980
	agggcactcg	cgctgccctt	ctagcatgaa	gtgtgtgggg	atttgcagaa	gcaacaggaa	2040
	acccatgcac	tgtgaatcta	ggattatttc	aaaacaaaag	tttacagaaa	catccaagga	2100
60	cagggctgaa	gtgcctccgg	gcaagggcag	ggcaggcacg	agtattttta	tttagctatt	2160
	ttattttatt	tacttacttt	ctgagacaga	gttatgtctt	tggtgcccag	gctggagtgc	2220
	agcggcatga	tcttggtcca	ctgcaacctc	cgtctcctgg	gttcaagcaa	ttctcgtgcc	2280
	tcagctcccc	aagtagctgg	gatttcaggc	gtgcaccacc	acaccgggct	aattttgtat	2340
	tttttagtaga	gatgggcttt	caccatgttg	gtcaagctga	tctcaaaatc	ctgacctcag	2400
65	gtgatccgcc	cacctcagcc	tcccaaagtg	ctgggattac	aggcatgagc	cactgcacct	2460
	ggcctattta	accattttta	aacttccctg	ggctcaagtc	acacccactg	gtaaggagtt	2520
	catggagttc	aatttccctc	ttactcagga	tttaccctcc	tttgatattt	tctgtaattt	2580
	ttcgtagact	ggggatacac	cgtctcttga	catattcaca	gtttctgtga	ccacctgtta	2640
	tccccatggga	cccactgcag	gggcagctgg	gaggctgcag	gcttcagggt	ccagtggggg	2700
70	tgccatctgc	cagtagaana	ctgatgtaga	atcagggcgc	aagtgtggac	actgtctcga	2760
	atctcaatgt	ctcagtgtgt	gctgaaacat	gtagaatata	aagtccatcc	ctcctactct	2820
	actgggattg	agccccctcc	ctatcccccc	ccaggggcag	aggagttcct	ctcactcctg	2880
	tgagggaagg	aatgatactt	tggtattttt	cactgctggg	actgaatcca	ctgtttcatt	2940
	tggttggtttg	tttggttttg	tttgagaggg	ggtttctact	ttgttgctca	ggctggaggg	3000
	agtgcgaatg	cgcgactctg	gcttactgca	gcctctgcct	cccagggttc	agtgaattct	3060
75	ctgcttccgc	ctcccatttg	gctgggatta	caggcacccg	ccaccatgcc	cagctaatatt	3120
	tttgatattt	tagtagagac	gggggtgggt	gggggttacc	atgttggtcc	ggctgggtct	3180
	gaacttctga	cctcagatga	tccacctgcc	tctgcccctt	aaagtgtctg	gattacaggt	3240

007260-942860

	gtgagccacc	atgcccagct	cagaattttac	tctgttttaga	aacatctggg	tctgaggtag	3300
	gaagctcacc	ccactcaagt	gttgtggtgt	tttaagccaa	tgatagaatt	tttttattgt	3360
	tgtagaaca	ctcttgatgt	tttactactgt	gatgactaag	acatcatcag	cttttcaaag	3420
	acacactaac	tgccaccata	atactgggggt	gtcttctggg	tatcagcaat	cttcattgaa	3480
5	tgccgggagg	cgtttctctg	ccatgcacat	gggtttaatt	actccagcat	aatcttctgc	3540
	ttccatttct	tctcttccct	cttttaaaat	tggtttttct	atgttggtct	ctctgcagag	3600
	aaccagtgtg	agctacaact	taacttttgt	tggaacaaat	tttccaaacc	gcccccttgc	3660
	cctagtggca	gagacaattc	acaaacacag	ccctttaaaa	aggcttaggg	atcataagg	3720
10	ggattttctag	aagagcgacc	tgtaatccta	agtatattaca	agacgaggct	aacctccagg	3780
	gagcgtgaca	gcccagggag	gggtgcgagg	ctgttcaaat	gctagctcca	taaataaagc	3840
	aatttctctc	ggcagtttct	gaaagtagga	aaggttacat	ttaaggttgc	gtttgttagc	3900
	atttcagtgt	ttgccgacct	cagctacagc	atccctgcaa	ggcctcggga	gaccagaag	3960
	tttctcgccc	ccttagatcc	aaacttgagc	aaccgggagt	ctggattcct	gggaagtcc	4020
	cagctgtcct	gcggttgtgc	cggggcccc	ggctctggag	ggaccagtgg	ccgtgtggct	4080
15	tctactgctg	ggctggaagt	cggcctcctc	agctctgcag	tccgaggcct	ggagcagggt	4140
	gcttggaacc	cgaggctgcc	ctccaccctg	tgccggcggg	atgtgaccag	atgttgccct	4200
	catctggcag	acagagtgcc	ggggccccag	gtcaaggccg	ttgtggctgg	tgtagggcgc	4260
	ccggtgcgcg	gccagcagga	gcgcctgggt	ccatttccca	ccctttctcg	acgggaccgc	4320
	cccgggtggg	gattaacaga	tttgggggtg	tttgctcatg	gtggggaccc	ctcgcccgct	4380
20	gagaacctgc	aaagagaaat	gacgggcctg	tgtaaggag	cccaagtcgc	ggggaagtgt	4440
	tgacgggagg	cactccggga	ggctcccggt	gcccgtccag	ggagcaatgc	gtccctgggt	4500
	tcgtcccccag	ccgcgtctac	gcgcctccgt	cctccccctc	acgtccggca	ttcgtgtgct	4560
	ccggagcccg	acgccccgcg	tccggacctg	gaggcagccc	tgggtctccg	gatcaggcca	4620
25	gcggccaaag	ggctgcgcga	cgcacctgtt	ccaggggcct	ccacatcatg	gccccctcct	4680
	cggtgtaccc	cacagcctag	gcccattcga	cctctctccg	ctggggccct	cgctggcgct	4740
	cctgcacctc	gggagcgcg	gcggcgcgcg	ggcggggaag	cgccggccag	acccccgggt	4800
	ccgcgcggag	cagctgcgct	ctcggggcca	ccagtggttc	ccagtggttc	cgccggcaca	4860
	gacgcccagg	accgcgctcc	ccacgtggcg	gagggaactg	ggaccggggc	acccgtcctg	4920
30	cccccttacc	ttccagctcc	gcctcctccg	cgccggaccc	gccccgtccc	gaccctctcc	4980
	gggtcccccg	cccagccccc	tccgggccc	ccagccccc	cccttctcct	tccgcccggc	5040
	cgccctctcc	tcgcccgcgc	agtttcaggc	agcgtgcgct	cctgctgcgc	acgtgggaag	5100
	ccctggcccc	ggccaccccc	gcgatg				5126
35	<210>	2					
	<211>	4042					
	<212>	DNA					
	<213>	Homo sapiens					
40	<400>	2					
	gtttcaggca	gcgctgcgct	ctgctgcgca	cgtaggggagc	cctggccccg	gccacccccg	60
	cgatgccgcg	cgtccccccg	tgccgagccg	tgccgtccct	gctgcgcagc	cactaccgcg	120
	aggtgctgccc	gctggccacg	ttcgtgccc	gctggggccc	ccagggtctg	cggtgtgtgc	180
	agcgcgggga	cccggcggtc	ttccgcgcgc	tggtggccca	gtgcctgtgt	tgccgtccct	240
45	gggacgcagc	gcccggcccc	gcggccccct	ccctccgcga	gggtgctctg	ctgaaggagc	300
	tggtggcccc	agtgctgcag	aggtctgtgc	agcgcggcgc	gaagaacgtg	ctggccttgc	360
	gcttcgcgct	gctggagggg	gcccgcgggg	gcccccccga	ggccttcacc	accagcgtgc	420
	gcagctacct	gcccacacag	gtgaccgacg	actgccccgg	gagcggggcg	tgggggctgc	480
	tgctgcgcgc	cgtaggggac	gacgtgctgt	ttcacctgct	ggcacgctgc	gcgctcttgc	540
50	tgctgggtggc	tcccagctgc	gcctaccagg	tgtaggggccc	ggcgtgttac	cagctcggcg	600
	ctgcccactca	ggcccggccc	ccgcacacag	ctagtggacc	ccgaaggcgt	ctgggatgag	660
	aacgggacctg	gaaccatagc	gtcaggggag	ccgggggtccc	cctgggacctg	ccagccccgg	720
	gtgagggagg	gcgcgggggg	agtgccagcc	gaagtctgccc	gttggcccaag	aggccccagg	780
	gtggcgctgc	ccctgagccg	gagcggagcg	ccgttgggca	ggggctcctg	gccccccggg	840
55	gcaggagcgc	tggaaccagt	gaccgtgggt	tcgtgtgtgt	gtcacctgccc	agaccctggc	900
	aagaagccac	ctcttttgag	gggtgcgctc	ctggcacgcg	ccactcccac	ccatccgtgg	960
	gcccgcagca	ccacgcgggc	cccccatcca	actcgcggcc	accacgtccc	tggaacacgc	1020
	cctgtccccc	gggtgtacgccc	gagaccaagc	acttctctca	ctcctcaggc	gacaaggagc	1080
	agctgcggcc	ctccttctca	ctcagctctc	tgaggccccag	cctgactggc	gctcggaggc	1140
60	tcgtggagac	catctttctg	ggttccaggc	cctggatgccc	agggactccc	cgaggttgc	1200
	cccgcctgccc	ccagcgctac	tggaataatgc	ggccccctgt	tctggagctg	cttgggaacc	1260
	acgcgcagtg	cccctacggg	gtgctcctca	agacgcactg	cccgtgcga	gctcgggtca	1320
	ccccagcagc	cggtgtctgt	gcccgggaga	agccccaggg	ctctgtggcg	gcccccgagg	1380
65	aggaggacac	agacccccgt	cgctgggtgc	agctgctccg	ccagcacagc	agccccctgg	1440
	aggtgtacgg	cttcgtgcgg	gcctgcctgc	gcccggctgt	gccccccagg	ctctggggct	1500
	ccaggcacaa	cgaacgcgcg	ttcctcagga	acaccaagaa	gttcatctcc	ctggggaagc	1560
	atgccaaagt	ctcgtgcag	gagctgacgt	ggaagatgag	gctgcgggac	tgccgttggc	1620
	tgccgaggag	cccagggtgt	ggctgtgttc	cgccgcgaga	gcaccgtctg	cgtgaggaga	1680
	tccttgcccaa	gttctctgac	tggtgtatga	gtgtgtacgt	cgtcagctgc	ctcaggtctt	1740
70	ttttttatgt	cacggagacc	acgttttcaa	agaacaggct	ctttttctac	cggaagagt	1800
	tctggagcaa	gttgcaaaag	attggaatca	gacagcactt	gaagagggtg	cagctcgggg	1860
	agctgtcggg	agcagaggtc	aggcagcatc	gggaagccag	gccccccctg	ctgacgtcca	1920
	gactccgctt	catccccaa	cctgacgggc	tgccggccgag	tgtagaacatg	gactacgtgc	1980
	tgaggagccag	aacgtttccg	agagaaaaga	gggcccagcg	tctcacctcg	agggtgaagg	2040
75	cactgttccag	cgtgtctaac	tacagagccg	cgccggcgccc	cgccctcctg	ggcgccctct	2100
	tgctgggccc	ggacgatata	cacagggcct	ggcgccacctt	cgtgctgcgt	gtgcccggcc	2160
	aggacccgccc	gcctgagctg	tactttgtca	aggtggatgt	gacgggcccgc	tacgacacca	2220
	tccccagga	caggctcacg	gaggtcatcg	ccagcatcat	caaaccccag	aacacgtact	2280
	gcgtgcgtcg	gtatgccgtg	gtccagaagg	ccgcccattg	gcacgtccgc	aaggccttca	2340

5	agagccacgt tgaggagac agggcagcag tcaggggcaa tgctctgcag acgggctgct cgaaaacctt tgcggaagac ttcagatgcc tggaggtgca tcaaccgcgg tgaagtgtca acatctacaa catttcatca cctccctctg gcgcgcggg tcaagctgac agacgcagct acccggcact aggccgagag ggcggccac aggctctgat aagggctgag caccgccagg tccatcccca accatccagg aggtgtgccc ttggggggag aaaaaaaaaa	ctctaccttg cagcccgcgt tggcctcttc gtcctacgtc cctgtgctac cctgcgtttg cctcaggacc agtgggtgaac ggccccacgg gagcgactac cttcaaggct cagcctgttt gatcctcctg gcaagtttgg ctactccatc cctcttgccc tcgacaccgt gagtcggaag gcctcagacc cagacaccag acccaggccc gtccggctga tgtccagcac ccagcttttc gattcgccat tggagaccct tgtacacagg gtgctgtggg aaaaaaaaaa	acagacctcc agggatgccg gacgtcttcc cagtgccagg ggcgacatgg gtggatgatt ctgggtccgag ttccctgtag ctattcccct tccagctatg gggaggaaca ctggatttgc ctgcaggcgt aagaacccca ctgaaagcca tccgaggcgg gtcacctacg ctcccgggga ttcaagacca cagccctgtc gcaccgctgg aggctgagtg acctgcgctc ctcaccaggga tgttccaccc gagaaggacc cgaggaccct agtaaaatac aa	agccgtacat tcgtcatcga tacgttccat ggatcccga agaacaagct tcttgttggg gtgtccctga aagacgaggc gggtcggcct cccgagacct tgctgcgcaa aggtgaacag acaggtttca catttttctc agaacgcagg tgacgtggct tgccactcct cgacgtgac tcctggactg acgcccgggt gagtcgtagg tccggctgag ttcacttccc ggccggcttc tcgcccctgc ctgggagctc gcacctggat tgaatatatg	gcgacagttc gcagagctcc gtgccaccac gggtcccatc gtttgccccg gacacctcac gtatggctgc cctgggtggc gctgctggat catcagagcc actctttggg cctccagacg cgatgtgtg gcgctgcatc gatgtcgctg gtgccaccaa ggggtcactc tgccctggag atggccaccc ctacgtccca cctgagtgag gcctgagcga cacaggctgg cactccccac ctcctttggc tgggaatttg gggggtccct agtttttcag	gtgggtcacc tccctgaatg gccgtgcga ctctccacgc attcggcggg ctcaccacag gtgggtgaact acggcttttg acccggaccc agtctcacct gtcttgcggc gtgtgcacca ctgcagctcc tctgacacgg ggggccaagg gcattcctgc aggacagccc gccgagacca gcccacagcc gggagggagg tgtttggcgc gtgtccagcc cgctcgggtc ataggaatag ttccaccccc gagtgaccaa gtgggtcaaa ttttgaaaaa	2400 2460 2520 2580 2640 2700 2760 2820 2880 2940 3000 3060 3120 3180 3240 3300 3360 3420 3480 3540 3600 3660 3720 3780 3840 3900 3960 4020 4042
30	<210> 3 <211> 11276 <212> DNA <213> Homo sapiens						
35	<400> 3						
40	acttgagccc tgggtgacaga tcttctcttg atacaaacac ttaaataagg acccacggta tcaaaaaagt yccaaggcgg aaccttgtcg cagctactcg gagccgggat aaaaataaaa aaaagcaaga cagaaataaa ttttgaaaag acctaataaa caaaggatca aaaatagata agcccaaaca agagaagccc gaattccaat tctacatggc aacaataaaa aaatcctcaa gtgatcaagt atgtgataca cagaaaaaag tatacaagaa aggccaaggt gagacctggg tagtcccagc tgacgtgagc gaataagaag aggaggtgga tttcaacata agccttccct tagtactaga ctggaagga gacttaagac caatgtacaa caaaaaagca tctacaatga	aagagttcaa atgagaccct ccacagtggg atgaaaatta aattgaaaaa taccagggta agaaaagcca gcagatcgcc ctactaaaaa ggaggctgag tgcgccattg aagttagaaa gcaaaactaa tgaaaactgaa ataaacaaaa ataaagtcag ctagaggcta aattcctaga gaccaataac aggacccaat cctactcaaa cagtattacc cagaaagaaa caaaacacta gggattttatt tcatcccaac atgtgataaa acatacaggc gggatgatgt ctacaaaaaa tagtctggag catgaacatg aaggagaagg ggagaagtgg ataaaaagccc ctaagatctg agtccctagc agaagtcaaa accactaaaa aaatcagtag gctacaaata aaactataaa	ggctacggtg gtctcaaaaa acaaaaccag aacaatatac tttatttaag gcagtgctaa ggcgacgtgg tgaggtcagg tacaataatt gcaggataac gactccagcc acttaaaaat cctaaaattg agataacaat ttgacaaacc agatgaaaaa ctatgagcaa tgcatacaac aataatggga ggcttccctg ctattctgaa ctgatttcaa gaaaactaca gcaaaacaaa ccagggatgg aaaatgaagt attctgcacc caggcacagt cttggggcca ctttttttaa gctgaggtgg tcaactgtact agaagggaga aagggggaag tatatgacag gaaaatgaca agagcaatca ttatcctgtt aactattaga tattttctata aaattaaaca atgttgataa	agccatgatt aaaaaaaaaa aaatcaacaa ttctgaatga caaatgataa gaagggaagt ctcatgcctg agttcgagac gtggggcatg cgcttgaacc tgggttaacaa acaacctaat gtaaaagaaa acaaaagatc tttggccaga agagacatta ctgtacacta ctaccaagat ttaaagccat ctggatttta aaatagagga aaccagacaa ggccaatatc ttaaacaaca aaggatgggt acaaaacta cttcatgata ggctcacacc ggagtttgag aaattagcca gagaatcact ccagcctaga agggaggagg ggaagggaag accgaggtag agggccactg gataagagaa tgcatgatgt gctgaaattt ttccaacagc gctaggaatt aagaaattga	gcaacaccac aattgaaata caagaggaaat ccagtgaatc cggaaacata tatagctata taatcccagc cagcctgacc gtggcacatg caggaggtgg gagtgaaacc gatgcacett agaaataata aacaaaatta ctaagaaaaa caactgatac ataaattgaa tgaacctaga aataaaaaag ccaatcattt aagaataact aaacacatca cctgatgaat ccttcgaaaag caacatatgc tatgattatt aaaaccctca tgcatcccca actagcctgg ggcatgatgg taagcctagg caacagaaca aagggaggag gaggaagaag tattatgagg ttcaccactg agaaataaaa atgatcttat ggtacagcag aaacaatctg aaccacagaa agagggcaca	acgcccagcct atataaagca tttgaaaact aatgaagaaa acctctcaaa agcagctaca actttgggag aacacagaga cctgtaattc agggttgcgg ctgtctcaag aaagaactag aagatcagag aaagtgtggt aggaaagagt aacactagaa agaaactcaa ctcctagcaa aaagaagaat ccaaactcat aaaacaaaaa actgatacaa atcattcatt aatcaatca tcactttatg aaaaccctgg gcactctggg gcaacaaaaa catatccctg aggctcaggc agacccact gaggagaagg aagaaacata aaaaactgaa gtattcaaca aaaaagaaac gtgaaagatc aaaaaagaaa	60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020 1080 1140 1200 1260 1320 1380 1440 1500 1560 1620 1680 1740 1800 1860 1920 1980 2040 2100 2160 2220 2280 2340 2400 2460 2520
75							

	agatatattcca	tgtttcataga	ttggaagaat	aaatactgtt	aaaaatgtcca	tactacccea	2580
	agcaatatttac	aaattccaatg	caatccctat	taaaatacta	atgacgttct	tcacagaaat	2640
	agaagaacaca	attctaatag	ttgtacagaa	ccacaaaaga	cccagaatag	ccaaagctat	2700
5	cctgacaaa	aagaacaaaa	ctggaagcat	cacattacct	gacttcaaat	tatactacaa	2760
	agctatagta	acccaaaact	catggtactg	gcataaaaac	agatgagaca	tggaccagag	2820
	gaacagataa	gagaattccag	aaacaaatcc	atgcattctac	agtgaactca	tttttgacaa	2880
	aggttgccaag	aacatacttt	ggggaaaaga	taatctcttc	aataaatggt	gctggaggaa	2940
	ctggatatcc	atatgcaaaa	taacaatact	agaactctgt	ctctcaccat	atacaaaagc	3000
10	aaatcaaaat	ggatgaaagg	cttaaattcta	aaacctcaaa	ctttgcaact	actaaaagaa	3060
	aacacccgag	aaactctcca	ggacatttga	gtgggcaaa	actctttgag	taattccctg	3120
	caggcacag	caaccaaaagc	aaaaacagac	aaatgggac	atatcaagtt	aaaaagcttc	3180
	tgccagcaat	aggaacaact	caacaaagag	aagagacaac	ccacagaatg	ggagaattata	3240
	tttgcacaa	attcatctaa	caaggaatta	ataaccagta	tataataagga	gctcaaaacta	3300
15	ctctataaga	aaaacaccta	ataagctgat	tttcaaaaat	aagcaaaaaga	tctgggtaga	3360
	cattttctcaa	aataagtcac	acaaattgga	aacaggcatc	tgaaaattgt	ctcaacacca	3420
	ctgatcatca	gagaaattgca	aactcaaaat	actatgagag	atcatctcat	ccaggttaaa	3480
	atggcttttta	ttcaaaagac	aggcaataac	aaatgccagt	gaggatgtgt	ataaaaggaa	3540
	acccttggac	actgttgggt	ggaattgaaa	ttgctaccac	tatgggaac	agtttgaag	3600
20	ttcttcaaaa	aactaaaaat	aaagctacca	tacagcaatc	ccattgtctag	gtatatactc	3660
	caaaaaagg	aactcagtgta	tcaacaagct	atctccactc	ccacatttac	tgcagcactg	3720
	ttcatagcag	ccaaggtttc	gaagcaacct	cagtgctccat	caacagacga	atggaaaaag	3780
	aaaatgtggt	gcacattatg	aatggagtac	tacgcagcca	taaaaaagaa	tgagatcctg	3840
	tcagttgcaa	cagcattggg	ggcactgttc	agtatgttaa	gtgaaataag	ccaggcacag	3900
25	aaagacaaac	ttttcatggt	ctcccttact	tgtgggagca	aaaaataaa	caattgacat	3960
	agaaattagag	gagaattggt	gttctagagg	gggtggggac	agggtgacta	gagtcacaaa	4020
	taattttattg	tatgttttaa	aataactaaa	agagataaat	tgggttgttt	gtaacacaaa	4080
	gaaaggataa	atgctttgaag	gtgacagata	ccccatttac	ctcgatgtga	ttattacaca	4140
	ttgtatgcct	gtatcaaaat	atctcatgta	tgctatagat	ataaacctta	cttatattaaa	4200
30	aattaaaaatt	ttaatggcca	ggcaccgttg	ctcatgtccg	taatcccagc	actttggggag	4260
	gccgaggcgt	gtggatcacc	tgaggtcagg	agtttgaaac	cagcttggcc	accatgatga	4320
	aacctgtctg	ctcataaaga	tacaaaaatt	agccaggcgt	ggtggccacat	acctgtagtc	4380
	ccaactactc	aggaggctga	gacaggagaa	ttgcttgaac	ctggggaggcg	gaggttgcag	4440
	tgagccgaga	tcatggccat	gcactgcagc	ctgggtgaca	gagcaagact	ccatctcaaa	4500
35	acaaaaacaa	aaaaaagaag	attaaaattg	taatttttat	gtaccgtata	aatatatact	4560
	ctactataatt	agaagttaaa	aattaaaaac	attataaaag	gtaattaaacc	acttaactta	4620
	aaataagaat	aalgttatgtg	cttctgaaga	cttctgaaga	agttaaagtt	atggccacga	4680
	tggcagaat	gtgaggaggag	aacagtgtaa	gttactgttg	ttagacgctc	atactctctg	4740
	taagtgtact	aatttttaacc	aaagacaggc	tgggagaagt	taaaaggcca	ttctataagc	4800
40	cctaaaaacaa	ctgcttaata	tggtgaaagg	taatctctat	taattaccac	ttaattacaga	4860
	tatcttctaaa	atcgagctgc	agaattggca	cgtctgatca	caccgtctct	tcattcacgg	4920
	tgcttttttt	cttgttggtc	tggagatttt	cgatttgtgt	ttcgtgtttg	gttaaaacta	4980
	atctgtatga	atcctgtaac	gaaaaatggt	gggtgatttc	tcagaagaa	ttagagtacc	5040
	tggcaggaag	cagggtggctc	tgtggacctg	agccacttca	atcttcaagg	gtctctggcc	5100
45	aagaccagag	tgcacaggca	aggcctgatg	acccgaggac	agggaaagctc	ggatgggaag	5160
	ggcgatagag	aagcctgcct	cgttggtgag	cagcgcatga	agtgcctcta	tttacgcttt	5220
	gcaaagattg	ctctggatac	catctggaaa	aggcgggccag	cgggaatgca	aggagtcaga	5280
	agcctctctg	tcaaaaccag	cgcagcagct	atggcgccca	cccgggcgtg	tgcagagggg	5340
	agaggagtga	aggcacctcg	aagtatggct	taaatctttt	tttcaactga	agcagtgacc	5400
50	aaggtgtatt	ctgaggggaag	cttgagtttag	gtgccttttt	taaaacagaa	agtcatggaa	5460
	gcacctttct	caaggggaaa	ccagacgccc	gctctgcggt	cattttacctc	tttctctctc	5520
	ccctctcttg	ccctgcgggt	ttctgatcgg	gacagagtga	cccccggtga	gcttctccga	5580
	gcccgctgctg	aggacctctc	tgcacagggc	ttccacagacc	cccgccttgg	agagaggagt	5640
	ctgagcctgg	cttaataaca	aactgggatg	tggctggggg	cggacagcga	cggcgggatt	5700
55	caaagactta	attctcatag	taaatccaag	ctttccacct	ccgaattgga	ttggatt	



5	ttccacaagc gccccacagc ccgacccccg ttaaacaac gaggaacatg tgccacactcc aggggagtg ccttttacta cataggggag cctgggcagg aaccgcgccg atccttcggg aggagggtca ggaggggagg gggacccctcc atcgtggacc gtgtgtgggg aaaaacaagg ggcaggcacg ggtatgctct cgtctcctgg gtgcaccacc gtcaagctga ctgggattac ggctcaagtc gttaccctcc catattcacac gaggtgcag atcagggcgc gtagaattta ccaggggcag cactgctggt ggtttcactc gctctgcctc caggcaccgc ggggttcacc tctgctctcc tctgtttaga ttaaagccaa gatgactaag gtcttctggg gggttgaatt tgtgttttct tggaacaaat ccctttaaaa agtattttaca ctgttcaaat aaggtttacat atccctgcaa aaaccggagt ggctctggag agctctgcag tgccggcggg gtcaaggccg ccattttccca tttgctcatg tgtcaaggag gcccgtccag cctcccttcc gagggacccc cccagggcct cctctctccg ggcggggaag ggccgggctc gagggactgg cgccggacccc cccagccctc agcgctgcgt	actaagcatc cctgggaatt ctgttttatt tggttaaaca ccgtttataa atgggatacg ttaggggggt aagccagttt tggggatggg ataatgctct gccccagggc actacctgca gaggggggca cctcgagccc acggagcctg tccggcctcc atttgcagaa tttacagaaa agtgtattta tgttgccag gtttcaagcaa acacccggct tctcaaaatc aggcatagac acacccactg tttgatattt gtttctgtga gcttcagctc aagtgtggac aagttccatc aggagtccct actgaatcca ttgttgctca cccagggtca ccaccatgcc atgttgcca aaagtgcctg aacatctggg tgatagaatt acatcatcag tatcagcaat actccagcat atgttggtct tttccaaacc aggcttaggg agacgaggct gctagctcca ttaagggtgc ggcctcgga ctggattcct ggaccagtgg tccgaggctt atgtgaccag ttgtggctgg ccctttctcg gtggggaccc cccaagtgcg ggagcaatgc acgtccggca tggggtctcg ccacatcatg ctggggccct cgcggcccg ccagtggaat ggacccgggc gccccgtccc ccccttctct cctgctgcgc	ctcttcccaa cacgtgacta ttaatagcta aacgggtcca agcctgcagg tacgcaacat taaggacggg cctggttctg ggaaccggga agagatgccc ctttgcagg ggcccgaaaa gcctcaggac agccctgcaa cagcagggaag gtgccatagg gcaacaggaa catccaagga tttagctatt gtctgagctg ttctcgtgcc aattttctat ctgacctcag cactgcacct gtaaggagtt tctgtaattc ccactgttta ccagtggggt actgtcctga ctcttactct ctcactcctg ctgtttcatt ggctggaggg agtgttctc cagctaattt ggctgtctc gattacaggt tctgaggtag tttttattgt cttttcaagg cttttcaagg cttccattga aatcttctgc ctctgcagag gcccctttgc atcactaagg aacctccagc taaataaagg gtttgttagc gacccagaag ccgtgtggct ggagccagg atgttggcct tgtgagcgcc acgggacccg ctcgccgcct ggggaagtgt gtcctcgggt ttcgtgtgtg gatcaggcca gcccctccct cgctggcgct accccgggt cgcgggcaca acccgtcctg gaccttccct tccgcggccc acgtgggaag	aagaccacgc cgcacatcat caaagcaggg tccgcacggg catctcaagg gctcaaaaag gggggcggca atgggtattg ggctgtgcca acgtcctgat gtgatctccg gtaatccagg gatggaggca gcgcctccag gcacggctgg agggcactcg acctatgcac cagggtgaa gtgcctccgg tatttttatt agcggcatga gtgatccgcc ggcctattta catggagttc ttcgtagact tcccatggga tgccatctgc atctcaatgt actgggattg tgagggaagg tgttggtttg cgcgatcttg ctgcttccgc tttgtatttt gaacttctga gtgaccacc gaagctcacc tgttagaaca acacactaac tgccgggagg aatttccctc ttccatttct aacagtgta cctagtggca ggatttctag gagcgtgaca aatctcctcc atttcagtgt tttctcgccc cagctgtcct catctgccag ccggtgcgcg cccgggtggg gagaacctgc tgacgggagg tcgtccccag ccggagcccg gcggccaaag cggtttaccc ctcgcacctc ccgcgccggg gacgcccagg cccttcacc gggtcccccg cgccctctcc cctgggcccc	ctggacattt gtacacactc aaatccctgc gaattacgct aaagaatttc gctgggggct ctcagttatg tctttgccat tccccaaaac tgaggaccct gaggtctggg ggttctggga gtcagttcga aagctggaaa cccttagccc cgctgcccct tgtgaatcta gtgcctccgg tacttacttt tcttggtcga cactcaacct aagtagctgg gatgggcttt cacctagtgt tcccaaagtg accattttta aatttccctc ggggatacac cccactgcag cagtagaaac ctcagtgtgt agcccccctc aatgatactt tttgttttgt cgcgatcttg ctcccatttg tagtagagac cctcagatga atgcccagct cactcaagt ctcttgatgt tgaccccata cgtttccctc tcttctccct agctacaact gagacaattc aagagcgacc gcccaggagg ggcagtttct ttgcccacct ccttagatcc gcggttgctc ggctgggaag cgaggctgcc acagagtgcc gccagcagga gattaacaga aaagagaaat cactccggga ccgcgtctac acgcccgcg ggctgcggca cacagcctag gggagcgcca cagctgcgct accgcgtccc ttccagctcc cccagccccc tcggggcgct gagttcaggc gcatg	7260 7320 7380 7440 7500 7560 7620 7680 7740 7800 7860 7920 7980 8040 8100 8160 8220 8280 8340 8400 8460 8520 8580 8640 8700 8760 8820 8880 8940 9000 9060 9120 9180 9240 9300 9360 9420 9480 9540 9600 9660 9720 9780 9840 9900 9960 10020 10080 10140 10200 10260 10320 10380 10440 10500 10560 10620 10680 10740 10800 10860 10920 10980 11040 11100 11160 11220 11276
70	<210> 4 <211> 104 <212> DNA <213> Homo sapiens					
75	<400> 4 gtggggcctcc catgcccagga	ccgggggtcgg gcagcgcagg	cgctccggctg cgactcaggg	gggttgaggg cgcttcccc	cgcccggggg gaaccagcga	
					60 104	

```
<210> 5
<211> 8616
<212> DNA
<213> Homo sapiens
```

5

[illegible][illegible]

5	ggctggagtg	taatggcaca	atctcggtct	actgcaacct	ctgctctctc	ggttcaagca	4380
	gttctcattc	ctcaacctca	tgagttagctg	ggattacagg	cgccccacc	cacgcctggc	4440
	taatttttgt	attttttagta	gagataggct	ttcaccatgt	tggccaggct	ggctctcaaa	4500
	tctcgacctc	aagtgtagctg	cccgctctgg	cctcccacag	tgtctgggatt	acagggtgcaa	4560
10	gccaccggtg	ccggcatacc	ttgatctttt	aaaatgaagt	ctgaaacatt	gctacccttg	4620
	tctctgagca	taagaccctt	agtgtatttt	agctctggcc	acccccagc	ctgtgtgctg	4680
	ttttccctgc	tgacttagtt	ctatctcagg	catcttgaca	cccccaaga	ctaagcatta	4740
	ttaattattgt	tttccgtggt	gagtgtttct	gtagctttgc	ccccgcctt	ctttttccct	4800
15	tttgttcccc	gtctgtcttc	tgtctcaggc	ccgcgctctg	gggtccctt	ccctgtctct	4860
	tgcgtgggtc	ttctgtcttg	ttatttgttg	taaacccag	ctttacctgt	gctggccttc	4920
	atggcatctg	gcgacgtccg	gggacctctg	cttatgatgc	acagatgaag	atgtggagac	4980
	tcacgaggag	ggcggtcatc	ttggccgctg	agtgtctgag	gcaccacgtg	gccagcgctc	5040
20	cttagccagt	gagtgacagc	aacgtccgct	cggcctgggt	tcagcctgga	aaacccagg	5100
	catgtcgggg	tctgggtggt	ccgcggtgtc	gagtttgaaa	tcgcgcacaa	ctgcgggtgt	5160
	gcgcagctc	tacgggtgct	gctctggcgg	ggagtgtctg	cttctccct	tctgcttggg	5220
	aaccaggaca	aaggatgagg	ctccgagcgc	ttgtgcacca	acaggagcat	gacgtgagcc	5280
25	atgtggataa	ttttaaaatt	tctaggtctg	gcgcggtggc	tcacgctgtg	aatccccaga	5340
	ctttggggag	ccaagggcgg	tggatcacga	ggctcaggag	tcgagaccat	cctggccaac	5400
	atgatgaaac	cccatctgta	ctaaaaaac	aaaaattagc	tggggctggt	ggcgggtgcc	5460
	tgtaatccca	gctactcggg	aggctgaggc	aggagaattg	cttgaaacct	ggaagtggaa	5520
30	gttgacgtga	gcgcacattg	caccactgca	ctcagcctg	gcaacacagc	gaagactctg	5580
	ctcaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aattctagta	gccacattaa	aaaagtaaaa	5640
	aagaaaaaggt	gaaattaatg	taataataga	ttttactgaa	gccccagcat	tcacacacct	5700
	atcatttttg	gggtgttatg	tgtgggagcat	cactcacagg	acatttgaca	ttttttgagc	5760
35	tttgtctcgc	tgactcccggt	tgtaggctcc	gtgcgtggcc	atctcggcct	ggacctgctg	5820
	ggcttcccc	ggccatggct	gttgtagacc	atgggtcagg	tccgggatga	ggctcgccagg	5880
	ccctcagtg	gctggatgtg	cagtgctcgg	atggtgcacg	tctgggatga	aggtcgccagg	5940
	ccctgctgtg	agctggatgt	gtggtgtctg	gatggtgcag	gtcaggggtg	aggtctccag	6000
40	ggcctcggtg	agctggaggt	atggaagtcg	gatgatgcag	gtccgggggtg	aggtcgccag	6060
	gcctctgctg	gagctggatg	tgtggtgtct	ggatggtgca	ggctcagggt	gaggtctcca	6120
	ggcctcggt	aagctggagg	tatggagctc	ggatgatgca	ggctcgggggt	gaggtcgcca	6180
	ggcctctgct	tgagctggat	gtggtgtgtc	tggatggtg	aggtctgggg	tgaggtcacc	6240
45	aggccctgcg	gtgagctggg	tgtgcggtgt	ctggatggct	caggtctgga	gtgaagtgcc	6300
	cagacgggtg	cagaccatgc	gggtgacgtg	atatgcggtg	tccggatggt	gcaggtctgg	6360
	ggtgaggttg	ccaggccctg	ctgtgagttg	gatgtggggt	gtccggatgc	tgacaggtccg	6420
	gtgtgaggtc	accaggccct	gctgtgagct	ggatgtgtgg	tgtctggatg	gtgcaggttc	6480
50	ggggtgaagg	tcgccaggcc	ctcgtctgtg	agctggatgt	gtggtgtctg	gatggtgcag	6540
	gtctggagtg	aggtcgccag	acccctcggt	agctggatgt	gcaggttcca	gatggtgcag	6600
	gtccgggggtg	aggtcgccag	ggcctcggt	gagctggatg	tgcggtgtct	ggatggtgca	6660
	ggtctggagt	gaggtcgcca	ggcctcggt	tgagctggat	tatggagttc	ggatggtgcc	6720
55	ggctcggggt	gaggtcgcca	gacctgtgct	gtgagctgga	gtgcggtgtc	tggatggtac	6780
	aggtctggag	tagggtgcc	agacctgtct	gtgagctgga	tatgcggtgt	ccgatgggtg	6840
	caggtcaggg	gtgaggtctc	caggccctcg	gtgagctgga	ggataggagt	ccgatgatg	6900
	caggtccggg	gtgaggtcgc	caggccctgc	tgtgaactgg	atgtgcggcg	tctggatggt	6960
60	gcaggtctcg	gggtgtggtc	ccaggccctc	ggtagctcgg	aggtatggag	tcgggatgat	7020
	gcaggtccgg	gggtgaggtc	ccaggccctg	ctgtgagctg	gatgtgcggc	gtctggatgg	7080
	tgcaggtctg	gggtgtggtc	gccaggccct	cggtagctgt	gaggtatgga	gtccggatga	7140
	tgcaggtccg	gggtgaggtt	ggcaggccct	gagctggatg	ggatgtgtct	tatccggatg	7200
65	gtgcagttcc	gggtgaggtc	gccaggccct	gctgtgagct	ggatgtgctg	tatccggatg	7260
	gtccaggtct	ggggtgaggt	caccaggccc	tgcggtgagc	tggttgtgctg	gtgtcccggt	7320
	gtccaggtc	cggggtgagt	tcgccaggcc	ctgggtgagc	tggtatgtgc	gtgtcccggt	7380
	gtccggatgg	tgcaggtcca	gggtgaggtc	gctaggccct	tgggtgggctg	gatgtgccgt	7440
70	gtccggatgg	tgcaggtctg	gggtgaggtc	gccaggccct	tggtagctgt	gatgtgcggt	7500
	gtctgcagtg	tgcaggtctg	gggtgaggtc	gccaggccct	tggtagctgt	gatgtgtggt	7560
	gtccggatgg	tgcaggtccg	gctgtgaggt				

```
<210> 6
<211> 2089
<212> DNA
<213> Homo sapiens
```

<400> 6  
 gtactgtatc cccacgccag gcctctgctt ctcgaagtc tggaaacacca gcccggcctc 60  
 agcatgcgcc tgtctccact tgctctgctt tccctggctg tgcagctctg ggctgggagc 120  
 caggggcccc gtcacaggcc tggtccaagt ggattctgtg caaggctctg actgcctgga 180  
 gctcacgttc tcttacttgt aaaatcagga gtttgtgcca agtgggtctct aggggtttgta 240  
 aagcagaagg gatttaaatt agatggaaac actaccacta gcctccttgc ctttccctgg 300  
 gatgtgggtc tgattctctc tctctttttt ttttcttttt tgagatggag tctcactctg 360  
 ttgccaggcc tggagtgcag tggcataatc ttggctcact gcaacctcca cctcctgggt 420  
 ttaagcgatt caccagcctc agcctcctaa gtagctggga ttacaggcac ctgccaccac 480  
 gcctggctaa tttttgtact tttaggagag acgggggttc accatgttgg ccaggctggt 540  
 ctcgaactca tgacctcagg tgatccaccc accttggcct cccaaagtgc tgggtttaca 600  
 ggctaagcca ccgtgcccag cccccgattc tcttttaatt catgctgttc tgtatgaatc 660  
 ttcaatctat tggatttagg tcatgagagg ataaaaatccc acccacttgg cgactcactg 720  
 cagggagcac ctgttcaggg agcacctggg gataggagag ttccaccatg agctaacttc 780  
 taggtggctg catttgaatg gctgtgagag tttgtctgca atgttcggct gatgagagtg 840  
 tgagattgtg acagattcaa gctggatttg catcagttag ggacgggagc gctggctctg 900  
 gagatgccag cctggctgag cccaggccat ggtattagct tctccgtgtc ccgccaggc 960  
 tgactgtgga gggctttagt cagaagatca gggcttcccc agctccccctg cacactcgag 1020  
 tccctggggg gccttgtgac accccatgcc ccaaactcagg atgtctgcag agggagctgg 1080  
 cagcagacct cgtcagagag aacacagctt ctgggctggg gaccccgacg tgggtgctgg 1140  
 gccatttcct tgcattctgg ggagggctcag ggccttccct gtgggaacaa gttaatacac 1200  
 aatgcacctt acctagactt tacacgtatt taatgggtg cgacccaaca tggctatttg 1260  
 accagtattt tggaaagaat ttaattgggg tgaccggaag gagcagacag acgtgggtgt 1320  
 ccccaagatg ctcttgttca ctactgggac tgttgttctg cctggggggc cttggaggcc 1380  
 cctctccctt ggacagggtt ccgtgccttt tctactctgc tgggcctgag gcctgcggtc 1440  
 agggcaccag ctccggagca ccgcggcccc cagtgtccac ggagtgccag gctgtgccc 1500  
 acagatgcgc aggtccaggt gtggccgctc cagccccctg gcccccatgg tgggttttgg 1560  
 gggaaaaggc caagggcaga ggtgtcagga gactgggtgg ctcatgagag ctgattctgc 1620  
 tccttggtg ctcttaccct ggggttccct agcagcctct cccgcctctt ccacttgaag ggatgtggct 1680  
 cttctacact ggggttccct cctggggcca ccttggggtt accccagtggt ctgtaccaga 1740  
 gggacaggca ctctgtgtgg aggggcatgg gttcacgtgg cccagatgac agcctgggac 1800  
 caggctccct ggtgctgatg gtgggacagt caccctgggg gttgaccgac ggactggcag 1860  
 tcccagggtt tgactatagg accaggtgtc cagggtccct gcaagtagag gggctctcag 1920  
 aggcgtcttg ctggcatggg tggacgtggc cccgggcatg gccttcagcg tgtgctgccc 1980  
 tgggtgcccc gagccctcac tgagtcggtg ggggtctgtg gcttccctg agcttcccc 2040  
 tagctctgtg tctggctgag caagcctcct gaggggctct ctattgcag 2089

<210> 7  
 211> 687  
 "<212> DNA  
 213> Homo sapiens

<400> 7  
 gtggctgtgc tttggtttaa cttccttttt aaacagaagt gcgtttgagc cccacatttg 60  
 gtatcagctt agatgaaggg cccggaggag gggccacggg acacagccag ggcatggca 120  
 cggcgccaac ccatttgtgc gcacagttag gtggccgagg tgccgggtgcc tccagaaaag 180  
 cagcgtgggg gtgtaggggg agctcctggg cagaggacag gctctgagga ccacaagaag 240  
 cagccggggc agggcctgga tgcagcacgg cccgaggttc tggatccgtg tcctgctgtg 300  
 gtgcgcagcc tccgtgagct tccgtttacg gggcccgggg accaggccac gactgccagg 360  
 agcccaccgg gctctgagga tcttggacct tgccccacgg ctctgacac ccacctgtt 420  
 ggctgagggt gctgcgggtg ccccgctcatc tgaggagagt gtgggggtgag gtggacagag 480  
 gtgtggcatg aggatcccggt gtgcaacaca catgcggcca ggaacccgtt tcaaacaggg 540  
 tctgaggaag ctgggagggt ttctaggttc cgggtctggg tggctgggga cactggggag 600  
 gggctgcttc tcccctgggt ccctatgggt ggggtgggac ttggccggat ccactttcct 660  
 gactgtctcc catgtgtgac ccgccag 687

<210> 8  
 <211> 494  
 <212> DNA  
 <213> Homo sapiens

<400> 8  
 gtgggtgccc gggacccccg tgagcagccc tgctggacct tgggagtggc tgctgatttg 60  
 gcacctcatg ttgggtggag gaggtactcc tgggtgggac gcagggagtg caggtgacct 120  
 tgtcactggt gaggacacac ctggcaccta ggggtgaggc cttcagcctt tctgacaga 180  
 catggggccc actgtgcacc ctgactgccc gggctcctat tccccaggag ggtccacttg 240  
 gattccagtt tccgtcagag aaggaaccgc aacggctcag ccaccaggcc ccggtgcctt 300  
 gcacccagct cctgagccag gggctcctc tctgaggct cagagagggg acacagccc 360  
 cctgccctt ggggtctgga gtgggtgggg tctagagag agtgggggac accgcaggc 420  
 caggccctga gggcagaggt gatgtctgag tttctgcgtg gccactgtca gtctcctgc 480  
 ctccactcac acag 494

<210> 9  
 <211> 865  
 <212> DNA  
 <213> Homo sapiens

&lt;400&gt; 9

5	gtaagggtcca	cgtgtgatag	tcgtgtccag	gatgtgtgtc	tctgggatat	gaatgtgtct	60
	agaatgcagt	cgtgtctgtg	atgcgtttct	gtgggtggagg	tacttccatg	atttacacat	120
	ctgtgatatg	cgtgtgtggc	acgtgtgtgt	cgtgggtgcat	gtatctgttg	cgtgcatatt	180
	tgtgtgtgtg	gtgtgtgtgg	cacgtgtgtg	tccatggtgt	gtgtgcctgt	ggtgtgcatg	240
	tgtgtgtgtc	tgtgacacgt	gcatgttcat	gctgtgtgct	gcatgtctgt	gatgtgccta	300
	tttgtgtgtg	gtgtgtgcat	gtgtccgtga	catatgcgtg	tctatggcat	gggtgtgtgt	360
	ggccccctgg	ccttactcct	tcttctccca	ggcatgggtcc	gcaccattgt	cctcacgtc	420
10	tcgggtgctg	gtttggggag	ctccacattc	agggctectca	cttctagcat	gggtgcccc	480
	gtcctgtcac	agggtctggc	cttggagact	gtaagccagg	tttgagagga	gagtagggat	540
	gctgggtgta	ccttctctga	ccccaggcac	ccagtcctggc	ctatgccggc	catgtcagg	600
	tccatgagat	ataggaaggg	tgattcaggc	ctcgtctccc	gggacacact	cctcccagag	660
	cgcccggggg	ccttggggct	cgccaggggg	gaaagggggc	ctgggcttgg	gttcccaccc	720
	agtgtgcatg	agcacgctgg	aggggttaagc	cctcaaagtc	gtgccaggcc	gggggtgcga	780
15	gggtgaagaag	tatccctgga	gcttcgggtc	ggggagaggc	acatgtggaa	acccacaagg	840
	acctctttct	ctgactttct	gagct				865

&lt;210&gt; 10

&lt;211&gt; 3782

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 10

25	tgtgggattg	gttttcatgt	gtgggatagg	tggggatctg	tgggattggg	ttttatgagt	60
	ggggtaaacac	agagttccaag	gcgagctttc	ttcctgtagt	gggtctgcag	gtgtcccaac	120
	agcttttattg	aggagaccat	atcttctctt	gaactatggg	cgggtttata	gtaagtccag	180
	gggtgtggagg	cctccccctg	gtctccctgt	ctgtttcttc	cactctgggg	tcgtgtgggt	240
	cctgtctgtgg	tgtgtggccg	gtgggcaggg	cttccaggcc	tccttgtgtt	catgtgctgt	300
30	gatgtggccc	tggctacgtg	ccgtccttgg	aattccccctg	cgagttggag	gctttcttct	360
	ttttcttttt	tctttctttt	tttttttttt	tgataacaga	gtctcgtctc	tttttgccca	420
	ggctggagtg	gtttggcggt	atcttggctc	actgcaacct	gtgtctcctg	agttcaagca	480
	attctcttgc	ctcagcctcc	caagttagctg	gaattatagg	cgccccacc	catgctgact	540
	aatttttgtg	atttttagtag	agacgaggtt	tctccatgtt	ggccaggctg	gtctcgaact	600
35	cctgacctca	ggtgatctct	ccacctcgcc	ctcccaaggt	gctgggatga	cagggtgtga	660
	ccgcgcgcgc	cgcccgagac	tcgcttctct	caagctccgt	gagatctgca	cgatagactg	720
	cctgacacct	tgggtgctgac	aacctccgtt	ttccttctcc	aggtctcgtc	aggggtcttt	780
	ccatttcatg	actctcttca	cagaagagtt	tcacgttgc	tgatttcccc	gctgtttctc	840
	cggttaattg	gtctctgtgt	ttatcgatgg	cctccttcca	tttcttttag	gctttgttta	900
40	ttgttgtttt	tccggctcct	tgaaggaaaa	gtttcgatta	tggatgtttg	aactttcttt	960
	tctaaacaag	catctgaagt	tgcctgtttc	cctctaaagc	agggatcccc	agggccccctg	1020
	ctgtggagtg	gcaccgctct	ggggcctgtt	ggaaacccgg	cgcacagcgg	gaggctaggt	1080
	gggggtgtggg	gagccagcgt	tcccgccctg	gccccgcccc	tctcagatca	gcagtgccat	1140
	gcggtgctca	gagggcgaca	caccctactg	agaactgtgc	gtgagagggg	tctagattct	1200
45	gtgctcctta	tgggaattcta	atgctctgat	atctgaggtg	gaaccgtttg	ctccaaaaac	1260
	catccccctc	cccactgctg	tcttgtggaa	aaatcgctct	ccacgaaacc	agtcctctgg	1320
	accacaatgg	ttgggggacc	tgtgtctaaag	acctgtctca	gcagcctctc	gtcagttgtg	1380
	atatattggc	tttctgtgtg	tgagtccaga	ataattacgg	atttctgtga	tgctttccgc	1440
	cgacctcaga	cccattgggt	atttgtgggc	gtgttgccctg	ctcctgggtt	gggaagggtg	1500
	caggccccat	gtaccttctc	gttactgcct	tccaggttgg	ttctcagggt	tgaatcgta	1560
50	tcgatgtggg	tttagccac	ggccctcgct	ccagctcctg	ggggctgggg	aacatgctga	1620
	agcacagagt	caccgtgcgc	gtcttttcat	gcctcacaag	ctcagggcct	cctgtgtccg	1680
	tgttagtggt	tgtcacgtgc	ctgctcacat	cctgtcttgg	ggacgcaggg	gcttagcagg	1740
	tcccgtagta	aatgacaagc	gtcctggggg	agtcctgcaga	ataggaggtg	gggggtgccg	1800
55	tctctctccc	gcgtcttccg	actcttctcc	tgctgtgtgt	gtggctgcac	ctgcacccct	1860
	gcaatccctc	cagcactggg	ctggagaggg	ccgggagctc	gagtgccact	tgtgcccact	1920
	gactgtggat	ggcagtcggt	cacgggggtc	tgatgtgtgg	tgactgtgga	tggcgggttg	1980
	tcacaggggt	ctgatgtgtg	gtgactgtgg	atggcggtcg	tggggtctga	tgtgggtctga	2040
	gtggatggcg	gtcgtggggg	ctgatgtgtg	gtgactgtgg	atggcggtcg	tggggctctga	2100
60	tgtgtgtgact	gtggatggcg	gtcgtggggg	ctgatgtggg	gactgtggat	ggcggtcgtg	2160
	gggtctgatg	tgggtgactgt	ggatggcagt	cgtgggggtc	gatgtgtggg	gactgtggat	2220
	ggcggtcgtg	gggtctgatg	tgggtgactgt	ggatggcagt	cgtgggggtc	gatgtgtggg	2280
	gactgtggat	ggcggtcgtg	gggtctgatg	tgtgtgactg	gtggatggcg	gtcgtggggg	2340
	ctgatgtgtg	gtgactgtgg	atggcggtcg	tggggtctga	tgtgtgggtga	ctgtggatgg	2400
	cggtcgtggg	gtctgatgtg	gtgactgtgg	atggcggtcg	tggggtctga	tgtgtgggtga	2460
65	ctgtggatgg	tgatcgggtc	caggggtctg	atgtgtgggt	actgtggatg	gcgggtcgtg	2520
	ggtctgatgt	gtgggtgactg	tggatgggtg	tcgggtcacag	gggtctgatg	tgtgggtggt	2580
	gtggatggcg	gtcgtggggg	ctgatgtgtg	gtgactgtgg	atggcggttg	gtcccggggg	2640
	tctgatgtgt	gggtgactgt	gatggcgatc	gggtcacagg	gtctgatgtg	tgggtgactg	2700
70	ggatggcggt	cgtgggggtc	gatgtgtggg	gactgtggat	ggcggtcgtg	gggtctgatg	2760
	tgtgtgtgact	gtggatggcg	gtcgtggggg	ctgatgtggg	gactgtggat	ggcggtcgtg	2820
	gggtctgatg	tgggtgactgt	ggatggcggt	cgtgggggtc	gatgtgtggg	gactgtggat	2880
	ggcggttggg	cccgggggtc	tgatgtgtgg	tgactgtgga	tggcggtcgt	gggggtctgat	2940
	gtgggtgactg	tggatggcag	tcgtgggggtc	tgatgtgtgg	tgactgtgga	tggcggtcgt	3000
75	gggtgtctgat	gtgtgggtgac	tgtggatggc	gggtcgtggg	tctgatgtgt	gggtgactgtg	3060
	gatggcggtc	gtgggggtcgtg	atgtgtgggt	actgtggatg	gcgggtcgtg	gggtctgatg	3120
	ggtgactgtg	gatggcggtc	gtgggggtcgtg	atgtgtgggt	actgtggatg	gtgatcgggtc	3180
	acaggggtctc	gatgtgtggg	gactgtggat	ggcggtcgtg	gggtctgatg	tgtgtgtgact	3240
	gtggatggcg	gtcgtggggg	ctgatgtggg	gactgtggat	ggcggtcgtg	gggtctgatg	3300

001260-9728560

5	tgtggtgact gtcacagggg actgtgtgat tgatgtgtgg tcgggtcacag actttgcgtc tggggttcat agtgtccagc ag	gtggatggcg tctgatgtgt gcgggtcggtg tgactgtgga gggtctgatg ctcggccccc cccgccatcg tctggccggg	gtcgtagggt gggtgactgt gggtctgatgt tggcgggtcgt tgtgttagct cggcccccgt ggcttggccg gcaggccaca	ctgatgtgtg gatggcggtc gtgggtgactg gggggtctgat gcaggtggag ttcccaaaca caggtccaca tttgtggctc	gtgactgtgg gtgggggtctg tggtatggcg gtggtgactg tcccagggtg gaagcttccc cgtcctgatc atgccctctc	atggcagtcg atgtgtgggtg tcgtgggggtc tggtatgggtg gtctgtagct agggcgtctc ggaagaaaca ctctgcccgc 3782	3360 3420 3480 3540 3600 3660 3720 3780
10	<210> 11 <211> 980 <212> DNA <213> Homo sapiens						
15	<400> 11 gtctggggcac aatcactggg tgtgatgggg tgccagacgt ccgcagtgcc caaaggaaag agctggcccc cgggcccgtg ggccctgtgg cctgttagca ctctgggcga ccagaccctg cgggtgggctg gtgtcacaca gatggccctg cggagggtct tctcccgctc	tgccctgcag ctcatgaccg gcgatgatgag gctgcattca tttgttcatg gtgtccccct tcagtgtgtg ttgagccacg ccctttgcag cttgcctggc atttccctgg tgcccgccag tgtgggtgtg ctctgcacata cattccagcc tggccacgtg gcttttcgca	gggtggggcac gacagactgt ctgtgtgcct ggcacctgct atttgctaaa ccttttaggag gtctgaggcc cccgcctgag atgtgggtctg tctaggggac ctcccagggt ctgggcagca agccccagctg gccccatgtg cagccccgca gtcctgtcctg	ggactcccag tggccctggg tggcgaatac cacgtttgac tgtcttctct ggcaggccat aaaggaaacg cgggcctctc tccacgtggc agtcgtgtcc gggggtggag actcctggat gacccacagg gtctgcagag cttcatcaca tctcagcacc	cagtgggtcc gggcagtggtg tgagctgggc tgccgggctc gccagttttg gtttgagccg tgtccccctt agtgtctggg cctgtggctc accgcatgag gtggcctggg cacatatgcc tggtcccagag actcggcccc aacactgacc caccggtcca	tccccctgggc gggaatgagc catgccaggc ctctccagtt atcttgaggc tgtcctgccc cttaggagga ctgtccacgt tttgcatgat gctcagagac ctgtctggac atccgggcca gagacgttct gccagcccac ccaaaaggga ctcccatgtg	60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 980
35	<210> 12 <211> 2485 <212> DNA <213> Homo sapiens						
40	<400> 12 gtgagtcagg tgctcacctc ttttctggcc ggctcggtct gggggtgtgga tgccgagcagc cggctctcac tctctgccct cccattctgga ccatggggga cgggtgtaga cacacctccc ggaggaaatt ggaggcctct tttatttaaa attataatat cacaaattgc taagcgcccc cgggcctcct gtggcagggc gggtagctgt tggtccagtt acagagagag gggctggccg gccccatcact ttttttgaga actcactgca gctgagatta gggtttttgc ctcggcctcc ctttttaagg gctgtcaagc tcgcgtggca actgtttgtc gtcactctga catgagctct catgagctct tcattccggg aaagaaaacc	tggccagggtg tctcctgccc cccggccccct gcggcagccg gttgcctcctg gtttgagcct acgcttgtat gtcgtgtgac aagtgcgggg ggcggcctgg gccacagtg ggcaggccatc cgtgcacact ctctgggatac aatataaacta ttattaaagt acatggcagc ccaggccccc tcgtgggtcgt tttgggggaat tctgtcctgt tggtccctga tttcccatcc gactcctaga gtgatatctg cggaaacgtca acctccgctc caggcaccca catgtttggc caaagtgtctg tgaccaccta tctcttagca gccatgcctt tgaaaaacgca aactaggggc tcaccgtgga tcaagtgtct ttgatgatcc	ccattgccct cttccccact ccggctcctg gagcggagca cgtggaggac gcagcttgtc ctctctctcc cccgcgaggg ttgaccgtgt gagagctgcc ctgggtgccac tgccctgcgac caaggctcatc gtctccagcg ttaattattg ataattagaa agagtgaatt agaattcgtc gaattttatt gtgaggtgat ccctaggaca ataaaaaactg catgtgtctca gttgggtgctg caccagcaag ctgtttgtctg cccgggttcc ccccctgcgc aggctggctc ggattacagg tagcgcttcc acaggagtgg ctgtgtgcac cccttggcat aaggttgtat caaattcctt gggtctgtga agagcaaggga	gcgggtgggt gnccttctgc ggctgcaggc ggtgccacac gaggggcggg agctccaagt cgatacaaaa gcgcgggctc agtttgcctc gtcacacagc atcacgtcct cctgtgtgtg agcaaggctca gataaaaggac cattataagt atattaagta ttggccgagg gacaaaagtc aagatggatc gactgcgtcc cggacaggcc cttcaaaaac caggggcgta gtgcttctgt gcaaaaagtg gttcttttct agtgaggctg agcatttctc ctggctaatt cgaactcctg tgtgagccat cgaaaaataac cgctcctgtg cttttaggtc ccttgttttg ccgttggcgc gaaaaaaaaa ataaactcta tgtggtcaca	ggggcgggctg ccggggccac tcccagggcc gagggcctgga gggtgtgtct tactactgac ggattttatc ttctctctgt tctcgggggg cactgggtga ctggatttta cctggggaga tccgcagtc tgtgcacagc aatcactaat gtacacacgt gacacgtgtg cctccccaga aagtcacgta tcatgcccctg cgaagctcta tggtgcccc tctgcttgcg gcaaaaagtg tttcttttct agtgaggctg ctgcctcagc tttgtatttt acctcagggtg cacgcccagc aggtccttgt gctctgggga cacgggggta agagtctctg gcagcgggta aaaggagctc agatttaaga cctgtggctg	gcagggccttc cagagctctcc ccggaaacat aatggcaagc gggtcagggtg gctggacacc cgatttctatc gactagattt cctgtgggtg gccacactca agtaaaaacca gtggtagcac ggtggaacgt ttcggaagct ggtatcagca tctggaaaaa cacatgtgtg gaagccacca ccgtccaggt acagacagga gtccccatcg aaaactaaga ttgactcgct cagtcctctt ttcttttttt ctcgattctca ctcccagagca tagtagagag atccacccac cggaaagcct tttgacgtat tggctgaggg ttctgtctct cttctcgttg catgtagggt gggttaagcat aaccttaagt gatctgtttc	60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020 1080 1140 1200 1260 1320 1380 1440 1500 1560 1620 1680 1740 1800 1860 1920 1980 2040 2100 2160 2220 2280

001260-542250

agccgcccca gtgcatgggt agagtgggga gcagggattt tttgttcaga ggtctcatct 2340  
 ggtagtgggt tgaggtgttt gccggtgaa tggtagacgt gtcgtttgtg tgtagtaggt 2400  
 tctgtgtctg tgtgtggctc ggtttgagtg tacgcatgtc cagcacatgc cctgcccgtc 2460  
 tctccactgt gtcttccgc cccag  
 <210> 13  
 <211> 1984  
 <212> DNA  
 <213> Homo sapiens  
 <400> 13  
 gtgaggcctc ctcttcccca ggggggcttg ggtgggggtt gatttgcttt tgatgcattc 60  
 agtggttaata ttcttgggtc tctggagacc atgactgtct tgtcttgagg aaccagacaa 120  
 ggttgacagc ccttcttggg atgaaggcgc accgggaagg ttgcacagcc tgaggactgc 180  
 gggctccacc caggtctctg ccagcggcca tgtccagagg cctcagggtt cagcaggcgg 240  
 gagggcgctc gggctgcatg atgagcatgt gaattcaaca ccgaggaagc acaccagctt 300  
 ctgtcaccgt acccaggttc cgttaggggtc ctgggggaga tggggctggt gcagcctgag 360  
 gccccacatc tcccagcagg cctctgcacg gtggcctgga ctgggcgctt cttcagccca 420  
 ttggccatcc cacttgcatt gggcttacac ccaaggacgc acacacctaa atatcggtgc 480  
 aaactaatgt ggttcaactc agctggcttt tattgacagc agttaccttt ttttttttaa 540  
 tactttaagt tctagggtag atgtgcacga cgtgcaggtt agttacatat gtatacatgt 600  
 gccatggttg tgtgtgtcac ccattaaact atcatttaca ttaggatat ctccaatgag 660  
 tatccctccc cactccccc atcccatgac aggccttggt gtgtgattgt cccaccctgt 720  
 tgtccaagtg ttctcattgt tcagtcccca cctgtgagtg agaacatgtg gtgtttggtt 780  
 tcttttcttt gcaatagtgt gctcagagtg atggtttcca gcttcgtcca tgtccctaca 840  
 aaggacatga actcatcctt tttttgact gcatagttat ccgtgtgta tatgtgccac 900  
 attttcttaa tccagttctat catcgatgga catttggtt ggttgcaagt ctttgcact 960  
 gtgaatagtg ccgaataaaa catacgtgtg catgtgtctt tatagacga tgatttataa 1020  
 tcctttgggt atataccag taatgggatg gctgggtcaa atggtatttc tagttctaga 1080  
 ccttgagaa atcacacac tgtcttccac aatgggtgaa ctagttaaca cttccacaa 1140  
 tgcttgaaaa gtgttctggt ctgggagagg atgtggacag cagttatttt tttatgaaaa 1200  
 tagtatcact gaacaagcag acagttagtg aaggatgcgt caggaaacct gcaggccaca 1260  
 cagccatttc tctcgaagac tccgggtttt tctgtgtcat ctttgaac acctgtcctt 1320  
 attatagcat gtacagtgga tccaggtttt tcttctaa ggttcaggt ctgagattga 1380  
 ataagtttat gtaacagaaa caaaaatttc ttgtacacac aactgtctct gggatttga 1440  
 gaaagatgtc ctcgagctg cggcacactg gtagggtcaca cagtgcacca tgcccagctt 1500  
 ccatggtcat ggggcgtgg gcttgggctt gctgaggacc acagctgcca tgctggtaaa 1560  
 cctgtggata ggatctgggt ctcggtatcat gctcagcccc agctttctta ccgtcttca 1620  
 gggcaccagc tggctcagag gggcgagggt gccgcgctg atggccttcg ttcgtcttca 1680  
 ttatttttcc ctaagagtct gagaagtggg aacactgagt acctataatg aatgaggaat 1740  
 gctggcacag aattgcacaa gctgtggtga ctggtggaaa gaaatttaag ttttctatt 1800  
 tgctgtagca gtaactgtga gagactcgt tcttattttt ggtgtgttaa attgtttgac 1860  
 aaccggtttg gagaattgta cttattttt agttaacctt gctgtgtatt ttccttatt 1920  
 tctagacag ttag  
 <210> 14  
 <211> 1871  
 <212> DNA  
 <213> Homo sapiens  
 <400> 14  
 gtgaggcccg cctccgtgtg ctgtggggac ctccacagcc tgtgggcttt gcagttgagc 60  
 ccccggtgtc ctgcccctgg caccgcagcg tttgtctctc caagtctct ctctctgccg 120  
 gtgctggatc cgcaagagca gaggcgcttg gccgtgcacc caggcctggg ggcgcagggg 180  
 cacttctggg agggagtggg tacctgtcag gcccttgctc tgcagagacc caccaggtt 240  
 acacacgtgg tgagtcaagg gcttgacctg gctcctgtct ctctttggaa agtcaagagt 300  
 ggcggctcct ggggccccag tgagaccccc aggagctgtg cacagggcct gcagggccga 360  
 ggcggcagcc tcttcccag ggtgcacctg agcctgcgga gagcaggagc tgcgtagtga 420  
 gctggcccac agcgttgcct cgggtcacgt tctgtcgtg ggttgtttgg gatcgggtgg 480  
 agaatttggg tttgtctagt gctgtgtct tgaaccacgg agatggctag gagtgggttt 540  
 cagagttgat ttttgtgaat caaactaaaa tcaggcacag gggacctggc ctcagcacag 600  
 gggattgttc aatgtgttcc cctcaaggg cgccccaca agcgggtggg ggtgttttaa 660  
 agtgcgattt gacgaggggac gagaacacct gaaagctgta aagggaaacc tcagaaaatg 720  
 tggcgccagc ggttggtttc aggtgctttg ctgggctgtg tttgtgaaaa cccatttga 780  
 ccgcgccctc aagtccaccc tccaggtcca cctccaggg ccgccttggg ctgggggttg 840  
 gcctggcggt ccttgtgccg cagcccggag cacagcaggg tgtgcacatt taaatccact 900  
 aagattcact cggggggagc ccaggtccca agcaactag ggtcagagg tcttgagggt 960  
 gctgaggggg gcttgcagag cgttctgtgt gcttctgtg ggcaagtcc tgagggtgct 1020  
 ggccaggggg gtggctcaga gtgtatgttg ggttcccacc gggggcagaa ctctgtctct 1080  
 gatgagtcgg cagccatgta acaggaaggg gtggccacag gggagctggg atgcaccagg 1140  
 ggaactgcgc agcttgcagg cctccaggg ccaggccaca ggaaggcgag ggggagccc 1200  
 ggggccacag cagaggccgc aggaagggaa ggggatgccc agggcagagc agaggctacc 1260  
 gggcacagga gggctcctgt agctgggtga cagctcagc tgcactcggg agggaaacct 1320  
 ctgcacgtga agctgacac tgggttgtcc cagctcacag cccagcagg tcccgcctc 1380  
 gagcaggaac tcagaaccct cccctttgtc taaagcacag cagatgcctt cagggcatt 1440  
 agggagaaaac gggcaaaagt gttgagaaac gttctaaaaa aaggtgggat ggttggaact 1500  
 tcttctccac attttagtct gctccqacc acagatgagt ctataacggg attgtggtg 1560

5	tgccatgggg gagtcctggc gggagacagg cctgccaggc aacaagggtg ctctgcctta	acacatgaga tgtccccggg gaaagcaccc ccagcaccc tcagggttacc g	tggaccatca ccaggccagg cgaagtcctg gtcctcaaat tcctgggtga	cagaggccac ttcttgcattg agcagggtctg accacttctc cggccccgca	tggggctgca ctcacctacc ggtccaggct tgggggtttc tcctggggct	cctcccatct tgtcctgccc cctcagagct caaagcattt gacattgccc	1620 1680 1740 1800 1860 1871
10	<210> 15 <211> 3801 <212> DNA <213> Homo sapiens						
15	<400> 15						
20	gtgagcgcac ccggttgcgtc ggccacaggg tgagggtgct cagacctggg gggacctgct cctctgaacc cagggccctt agtctacagg tgtggggggg gccccgtctc ctgtttcttt ataatcccag ccaacctaac cctggtggca aaccaggag acagagttag ggacaggtgt gaactggggg tgggtgttaa tggactttgc tggacacctc ggtgcagaca ggatgccggt ctggccctca cgctctctcc cttattttgc gcacagcatc gttctctcta catcagatgt gtgtgcttgc gggtgaactc gaagaaaaca ttcttgtcca gatggacaga gtatgtggca gactggaagc aggagacaca aggtagaacgt tgaggcaacg atggccattc cacattcctc aggggagcag ctcctgtctc agggtgagcc gatgcacact gagaccatc atgctggctc acttttcttg taattccact ccaggcaggg tattatgcat gtttaatggc cacaccccag gggcagttag ctggagacac gctcttccat ccacaaaaaa acagtttatt tgcaaaaaca gagttttggt cagcccccct aggctctttg acattccccc	ctggccggaa cacctctgct tgccccctgt cacaacggga tgactgagg gggctgagtg cgagacctg ttgggctgga atgccatgag gtctctacaa aggctcagac tatgaataaa cactttggga caacatagtg cacgctgta gcagaggttg acttcatctt ttttttattc tgccctcttc accagaggtt ctctttccag cgtgatgggg gggtcagaca ggatgccggt ctggcctcca cgctctctcc cttattttgc gcacagcatc gttctctcta catcagatgt gtgtgcttgc gggtgaactc gaagaaaaca ttcttgtcca gatggacaga gtatgtggca gactggaagc aggagacaca aggtagaacgt tgaggcaacg atggccattc cacattcctc aggggagcag ctcctgtctc agggtgagcc gatgcacact gagaccatc atgctggctc acttttcttg taattccact ccaggcaggg tattatgcat gtttaatggc cacaccccag gggcagttag ctggagacac gctcttccat ccacaaaaaa acagtttatt tgcaaaaaca gagttttggt cagcccccct aggctctttg acattccccc	gtggagcctg tccgtgtggg cccactctgg gcagttttct tgtcttcaga ctctcaaac gggacctgct gtctctccgc ttcatgatca aattctgggg acaaatgaat aagtatcaac ggccgaggtg aaattccatt gtccccgcta cagtgagccg aaaaaaaagt tgtcttcaga tgaaaggcac taaactgggg aatgctccct gagcagcagg tggtgccag ccccacagtc ctgcatgatt gcagtgctgg atgtattttt attgaaggac aaagccacag ccagaatatt ctaaaagctc tgtctgaagt ttaagaaaag tcccaaaaca aaaacggaag aaaagagagt tgtctttaca accagcaaca gggtgttggg tcaactcaga ttgtgcagct gttctcctaa tcacccagct ctctgcccga gtcgtgttgg ctaccagcag acgcagctga cttgccaaga gtttgcattg ccagacatta cagcaagtca gctctgccat tatttcaatg tgaatgtcat gcccctggag acgtgactc caaacacagt acctgtgttc ggctgagtta aggtttggat tggatggcat gcatgcccc tgcaggaggg g	tgccccgctg gcaggcgact gctgagcaga gtgctatttt aagcagttct cgaacacagg gggctgagtg tgtagccccc cgtgtgaccc tcttgtttcc tgaagatgga attccaggca ggtggatcac tctacttaaa tgccggaggc agatcacacc aaaaaagtat taatatttac accttcatgg tccgtctgtt gggttttgc tgacagcgc catgtccctg cctgcttccc tccacatttc ccataccagt taggacaggc aaaggacaga aggctagtgc ctgtgctccc agcagtggag atacagcaga tgaaaaagga cagtcagat ccctatctct gtgtgtgtaa gaaataaaac aaggacacac gaaactcagc gatttattta aggcgccctg ggtagaggag atgcatgatt caagtcagac gaaagaagaa atgcatgtga gagacctgtc ggttgaggca aatgtcctgt tatttaccat caaatacagg tcaaagaatt caaagctgga ctttggacat gatgcctcca tgtttagctg acgcccact cgccccagg tcacaggggc tactagcca gagtcctagg ggaagcggga ggggcaggca	tgctgcaggg aagggtcaga tctgtgggag aatggtgcac caagacgccc ggcatgagtc gctcatccac agggccatgg agagctcaag gaaatctgtg gctcacacct gagtttgagg attagcctgg gaatcatttg agcctgggca aaacctagtg ctagaggccg aagtggtgaa agtcacagtc gcagcaggtg gggagtgcca ctcccacaa ttacctggtc agcacctctt ctgtccactg ccagcctctg aggaaaatgg ggcatcaggt tggtcagagt gccatactca gcatctggga gatgggaatt ggtcagaact tggttaattg tgagaaaact aagggaaggg atgaaaccag cacagtgaaa tgagggtcctg gaaaggctcc gcagcctggc ccataggctc atggacgtct aactgacagc cccctccctc agctggaaag gtcttcccag ttccagtgtt gctaaggaga tttgaagaat tgtaaaagaa ggacatacat tcttcccct gtgccacctg cagtggtctc ccagggtctc ccatagggtc agatgagtag ggtcctgggt agtaggcacc ggaagcggag cctgtgtctg	60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020 1080 1140 1200 1260 1320 1380 1440 1500 1560 1620 1680 1740 1800 1860 1920 1980 2040 2100 2160 2220 2280 2340 2400 2460 2520 2580 2640 2700 2760 2820 2880 2940 3000 3060 3120 3180 3240 3300 3360 3420 3480 3540 3600 3660 3720 3780 3801	

001260 9428560



<210> 16  
 <211> 880  
 <212> DNA  
 <213> Homo sapiens

5

<400> 16  
 gtgagcaggc tgatgggtcag cacagagttc agagtttcagg aggtgtgtg gcaagtatgt 60  
 gtgtgtgtgt gtgcgcgcgt gcctgcaagg ctgatgggtga ctggctgcac gtaagagtgc 120  
 acatgtacgc atatacacgt gagcacatac atgtgtgcat gtgtgtacat gaaggcatgg 180  
 cagtgtgtgc acaggtgtgc aagggcacaa gtgtgtgcac atgcgaatgc acacctgaca 240  
 tgcatgtgtg ttcgtgcaca gtcgtgtggg cattcacgtg aggtgcatgc gtgtgggtgt 300  
 gcagtgtgag tagcatgtgt gcacataaca tgtattgagg ggtcctcgtg ttcaccccg 360  
 taggtcctca gcaccagtgc cactccttac aggatgagac ggggtcccg gccttggtg 420  
 gctgaggctc tgaagctgca gccctgaggg cattgtccca tctgggcatc cgcgtccact 480  
 cctctccttg tgggcttctg tgtccactcc cctctcctg tgggcattta catccactcc 540  
 actcctctct tctctgtggc atccgcgtcc actccctctc tctgtgggca tctgcgtcca 600  
 cctccctctc ctgtgggcat ttgcgtccac tccctctcct ggttccctcc tgtcttggcc 660  
 gagcctcggg ggcaggcaga tgacacagag tcttgactcg cccagggtgg ttcgcagctg 720  
 ccgggtgagg gccaggccgg atttacttgg gaagagggat agtttcttgt caaaatgttc 780  
 ctcttctctg ttccatctga atggatgata aagcaaaaag taaaaactta aaatcccaga 840  
 gaggtttcta ccgttttcta ctcttttctg gcgactctag 880

<210> 17  
 <211> 3186  
 <212> DNA  
 <213> Homo sapiens

25

<400> 17  
 gtgagccgcc accaaggggt gcaggcccag cctccaggga cctccgcgc tctgtccacc 60  
 tctgaccggg ggcttcacct tggaaactcct ggggttttagg ggcaaggaat gtcttacgtt 120  
 ttcagtgggt ctgctgcctg tgcacagtgc tggttcgcgtg gctctgtgca aagcacctgt 180  
 tctccatctc tgggtagttg taggagccgg tgtggcccca ggtgtcccca ctgtgcctgt 240  
 gcaactggcg tgggacgtca tggaggccat cccaggggcag caggggcatg gggtaaagag 300  
 atgtttatgg ggagtccttg cagaggaggg tgggaaaggtg tctgaacagt agatgggaga 360  
 tcagatgccc ggaggatttg ggggtcagc aaagagggcc gaggtgggtg caggtagggg 420  
 tgcgtggccc caccctcggg aaggtgcagc agagctgtgg ctccccacac agcccgcca 480  
 gcacctgtgc tctgggcatg gctgtgtctc tggaaacttc cctgtcctgg ctggtcaggg 540  
 ggtgccccct ccaagaatcg acaactttat cacagagggg agggccaatc tgtggaggcc 600  
 acagggccag cttctgcctg gagtcagggc aggtgggtggc acaagcctcg gggctgtacc 660  
 aaagggcagc cgggcaccac agggccgggc ctccacctca acaggcctcc cgagccactg 720  
 ggagtcgaat gccagggaat cgaagccctc gcccatgag ggctgagaag gagtgtgag 780  
 atttgtgtta cccaggggcg aggtgcgcg aattaccgtg cacacttgat gtgaaatgag 840  
 gtctgtcgtc atcgtggaaa cccagcaagg gctcacggga gatttttcca ttacaaggtc 900  
 gtaccatgaa aatggttttt aaccgagtg cttgcgcctt catgctctgg caggaggggc 960  
 agagccacag ctgcatgtta ccgcctttgc accagctcca gaggtttggg accaggctgt 1020  
 ctacgttcca ggggtgcgtc ggctcagacc gccctcctct ctgccttctc tctctgcctc 1080  
 aaatcttccc tctgttggag cctccttgag cgtgcctggg ccctcgtgca agctgtgacc 1140  
 ctcttttccc gaaaccttgg ggggtgtgctg gatacaggtg cactgagga ctggagggtg 1200  
 ctgacactgt ggttgacccc aggggtccag tggcgtgctt ggggctcctc tggggcatga 1260  
 tgaggtcaga ggagttttcc caggtaaaaa cctctgggaa actcccaggg ccattgtacc 1320  
 tgccacctgc tctcccata ttcagctcag tcttgtcctc atttcccac cagggtctct 1380  
 agctccgagg agctcccgtg gaggccctgg gctcagggca gggcggtga gtttcccac 1440  
 ccattgtggg acccttgggt agtcgcttga ttgggtagcc ctgaggaggc cgagatgca 1500  
 tggggccagg gccgtttcca aacacagagt caggcacgtg gaaggcccag gaatcccctt 1560  
 cctcggaggc aggagtgagg gaacggagag ctgggccccg atttccaggc agccaggctg 1620  
 cagtggggcg gctgtgggtg gtccacgtgg cgtggggggc ggggtctgat tcaaatccgc 1680  
 tggggctcgg ccttccctgg ccgtgtggc cgcgcctcca cacgggcttg ggggtggacg 1740  
 cccgacctct agcagggtggc tatttctccc tttggaagag agccctcac ccattgtatt 1800  
 tgttccctc ctgggtcagg agcgtggcgg tgtggcaacc ccgggacctt aggttattt 1860  
 atttgtttta aaacattctg ggcctggcct ccgttgttgc taaatgggga aaagacatcc 1920  
 cacctcagca gaggttactg gaggctgaaa ccgggtgtgt ggcttgactg gtgtgatctc 1980  
 aggtcattcc agaagtggct caggaagtca gtgagaccag gtacatgggg ggctcaggca 2040  
 gtgggtgaga tgaggtagac ggggggctca ggcagtgggt gaggccaggt acatgggggg 2100  
 ctacggcact gggtagatg aggtacacgg ggggctcagg cagaggggtca gaccaggtag 2160  
 acgggggctc tgatcacacg cacatgtgca catgtgtgt ttcatgttag 2220  
 ccagggtctg gcacacctgc cccaaagtcc caggaaagct agaggccaaa gatggaggct 2280  
 gacagggctg gcgcgggtggc tcacacctgt agtcccagca ctttgggagg ccgaggcgag 2340  
 aggatccctt gagcccagga gtttaagacc agcctgagca acatagtaga acccatctc 2400  
 tatgaaaaat aaaaacaaaa attagctgaa catggtgggtg tgcgcctgta gttccaatac 2460  
 ttggggaggct gaagtgggag gatcacttga gccaggaggg tggaaagctg agtgagctga 2520  
 gattgcacca ctgtactgca gcttgggtga cagagtggga gcccacttca acaacaacaa 2580  
 agaagactga caaatgcagt ttcttggaaa gaaacattta gtaggaactt aacctacaca 2640  
 cagaagccaa gtcggtgtct cgggtgcagt gcgatgagat gatgggtcct cacaccatca 2700  
 ccccagacc agggtttatg caccacaggg cagggtggct cagaagggat gcgcaggacg 2760  
 ttgatatacg atgacatcaa ggttgtctga cgaaggggcag gattcatgat aagtacctgc 2820  
 ttggtacaaa ggaacaatgg ataaacttga aaccttagag gccttcccg aacaggggtc 2880  
 aatcagaagc cagcatgggg ggttggcatc caggatggag ctgcttcagc ctccacatgc 2940  
 gtgttcatac agatgggtga cagaaacgca gtgtacctgt gcacacacag acacgcagct 3000

007260-34225560

	actcgcacac	acaagcacac	acacagacat	gcattgcattgc	atccgtgtgt	gtgcacctgt	3060
	gcccatgagg	aaacccatgc	atgtgcattc	atgcacgcac	acaggcaccg	gtgggcccac	3120
	gccacacccc	acgagcaccg	tctgattagg	aggcctttcc	tctgacgctg	tccgccatcc	3180
5	tctcag						3186
	<210> 18						
	<211> 781						
	<212> DNA						
10	<213> Homo sapiens						
	<400> 18						
	gtatgtgcag	gtgcctggcc	tcagtggcag	cagtgcctgc	ctgctgggtgt	tagtgtgtca	60
	ggagactgag	tgaatctggg	cttaggaagt	tcttaccctt	tttcgcatca	ggaagtgggt	120
15	taaccaaac	actgtcaggc	tcgtctgccc	gccctctcgt	ggggtgagca	gagcacctga	180
	tggaaaggac	aggagctgtc	tgggagctgc	catccttccc	accttgctct	gcctggggaa	240
	gcgctggggg	gcctgggtctc	tctgttttgc	cccatgggtg	gatttggggg	gcctggcctc	300
	tctgttttgc	cctgtgggtg	gattgggctg	tctccgctcc	atggcactta	gggcccctgt	360
	gcaaaccag	gccaagggtc	taggaggagg	ccaggcccag	gctacccac	ccctctcagg	420
	agcagaggcc	gcgtatcacc	acgacagagc	cccgcgctg	cctctgcttc	ccagtcaccg	480
20	tcctctgccc	ctggacactt	tgtccagcat	caggagggtt	tctgatccgt	ctgaaattca	540
	agccatgtcg	aacctgcggt	cctgagctta	acagcttcta	ctttctgttc	tttctgtgtt	600
	gtggaaattt	cacctggaga	agccgaagaa	aacatttctg	tcgtgactcc	tgcgggtgctt	660
25	gggtcgggac	agccagagat	ggagccaccc	cgcagaccgt	cgggtgtggg	cagctttccg	720
	gtgtctcctg	ggaggggagc	tgggctgggc	ctgtgactcc	tcagcctctg	ttttccccc	780
	g						781
	<210> 19						
	<211> 536						
	<212> DNA						
30	<213> Homo sapiens						
	<400> 19						
35	gcaagtgtgg	gtggaggcca	gtgcggggcc	cacctgccc	ggggtcatcc	ttgaacgccc	60
	tgtgtggggc	gagcagcctc	agatgctgct	gaagtgcaga	cgcccccg	cctgacctg	120
	ggggcctgga	gccacgctgg	cagccctatg	tgattaaacg	ctggtgtccc	caggccacgg	180
	agcctggcag	ggctccccaac	ttcttgaacc	cctgcttccc	atctcagggg	cgatggctcc	240
	ccacgcttgg	gagccttctg	acccttgacc	tgtgtctctc	cacagcctct	tccttgctcg	300
40	ctgccctgag	ctcctggggg	cctgagcaag	ttcttctccc	gccccgccc	tccagcgtca	360
	ctgggctgcc	tgtctgtctg	ccccggtgga	ggggtgtctg	tcccttcaact	gaggttccca	420
	ccagccaggg	ccacgaggtg	caggccctgc	ctgcccggcc	acccacacgt	cctaggaggg	480
	ttggaggatg	ccacctcttg	cctcttcttg	aacggagtct	gattttggcc	ccgcag	536
	<210> 20						
45	<211> 3179						
	<212> DNA						
	<213> Homo sapiens						
	<400> 20						
50	atctcatgtt	tgaatcctaa	tgtgactgct	atagacacca	ctgtatgcaa	ttacagaagc	60
	ctgtgagtga	acgggggtgt	ggtcagtgct	ggcccatggc	ctggtgtgtc	atttacggaa	120
	gtctatgagt	gaatgggtgt	gtggtcagtg	cgggcccatt	gcctggctgg	gcctggggag	180
	ttctctgagc	tgtgaggcag	gagggggaag	agggtagggg	atagacagt	ggagccccc	240
55	ccctggaaga	cataacagta	agtccaggcc	cgaagggcag	cagggatgct	gggggcccag	300
	cttggggcgg	ggggatgatg	gagggccttg	ccagggtggc	agggatgatg	ggggcccacg	360
	ctgggggtgg	aggggtgatg	gggggggctg	gtctgggtgg	cggggaagat	ggggaagcct	420
	ggctggggcc	cctcctcccc	tgcctcccac	ctgcagccgt	ggatccggat	gtgcttccct	480
	ggtgcacatc	ctctggggca	tcagctttca	tggaggtggg	gggcaggggc	atgacaccat	540
	cctgtataaa	atccaggatt	cctcctcctg	aacgccccaa	ctcaggttga	aagtcacatt	600
60	ccgcctcttg	ccatttctct	aagagtagac	caggattctg	atctctgaag	ggtgggtagg	660
	gtggggcagt	ggaggggtgt	gacacaggag	gcttcagggt	ggggctgggt	atgctctctc	720
	atcctcttat	catctcccag	tctcatctct	catctcttta	tcacttccca	gtctcatctg	780
	tcttctctct	atctcccagt	ctcatctgtc	atcctcttac	catctcccag	tctcatctct	840
	tatctcttta	tctcttagtc	tcattccagc	ttacttccca	gggctgggtg	caggctcgca	900
65	gtggagctgg	acatacgtcc	ttcctcaggg	agaaggaaat	ggaaggattg	cagaaacaca	960
	gagggggcgg	tcagagggac	gcagctcttg	ggtgaagaaa	cagccccctc	tcagaagttg	1020
	gcttggggca	cacgaaacgg	agggccctgc	gtgagtggtc	ccagagcctt	ccagcaggtc	1080
	cctgggtggg	ccttatggta	tggccgggtc	ctactgagtg	caccttggac	agggcttctg	1140
70	gtttgagtgc	agcccggacg	tgcctgggtg	cggggtgggg	gcttatggcc	actggatatg	1200
	gcgtcattta	ttgctgtctg	ttcagagaat	gtctgagtga	ccgagcctaa	tgtgtatggt	1260
	gggcccgaag	ccacagactg	tgtcgtaaat	gcactctggt	gcctggagcc	cccgtatagg	1320
	agctgtgagg	aaggaggggc	tcttggcagc	cggcctgggg	gcgcctttgc	cctgcaaaact	1380
	ggaaggggag	ggccccgggc	gcccgtgggc	cagcagctca	agtgagaggt	tggacagaac	1440
	agggcggggg	cttcccaggg	gcagagcctg	gtctcaggc	acacctgggt	ttgaataaca	1500
75	gacccaacag	tcaggccatt	gttcagctat	ccatcttcta	caaagctcca	gattcctgtt	1560
	tctccgggtg	ttttttgttg	aaattttact	caggattact	tatatTTTTT	gctaaagtat	1620
	tagaccctta	aaaaaggtat	ttgctttgat	atggcttaac	tcactaaagc	ctactcttat	1680
	ttgtctgttt	ttattttatta	ttattattat	tattagagat	ggtgtctact	ctgtcaccca	1740

007260-37225560

	ggttgttagt	gcagtggcac	agtcattggct	cgctgtagcc	gcaaaccccc	aggctcaagt	1800
	gatcctccgg	cctcagcttc	ccagagtgtc	gggattacag	gtgtgagcca	ctgcccttgc	1860
	ctggcacttt	taaaaaccac	tatgtaaggt	cagggtccagt	ggcttccaca	cctgtcatcc	1920
5	cagtagtttg	ggaagccgag	gcagaaggat	tgtctgaggc	caggagtttg	agaccagcat	1980
	gggtaacata	gggagacccc	atctctacaa	aaaatgcaaa	aagttatccg	ggcgtggggt	2040
	ccagcatctg	tagtcccagc	tgctcgggag	gctgagtggg	aggatcgctt	gagcccggga	2100
	ggtcatggct	gcagtgtgct	gtgattgtac	catcgactc	cagcctgggc	aacagagtga	2160
	gacctgtct	caaaaaaaaa	aaaaaaaaaa	gaaggagaag	gagaagagaa	gaagaaggaa	2220
10	gaaggaaaga	gaagaagaag	gaagaaggaa	gaaagaagga	gaaggaggcc	tgctaggtgc	2280
	taggtagact	gtcaaattct	agagcaaaat	gaaaataaca	aagttttaaa	gggaaagaaa	2340
	aaccccagct	ctttggactt	ccttaggcct	gaacttcata	tcaagcagct	tccttccaca	2400
	gacaagcgtg	tatggagcga	gtgagttcaa	agcagaaagg	gaggagaagc	aggcaagggt	2460
	ggaggctgtg	ggtgacacca	gccaggaccc	ctgaaaggga	gtgggtgttt	tcctgcctca	2520
15	gccccacgct	cctgccggtc	ctgcacctgc	tgtaaccgtc	gatgttggtg	ccaggtgccc	2580
	acctgggaag	gatgctgtgc	agggggcttg	ccaaactttg	gtgggtttca	gaagccccag	2640
	gcacttgtgg	caggcacaaat	tacagccctt	ccccaaagat	gccccacgtc	ttctcctgga	2700
	acctgtgaat	gtgtcacccg	caaggcagag	gctggtgaag	gctgcagggt	gaatcacggc	2760
	tgccagtcag	ccgatcttaa	ggtcatcctg	gattatctgg	tgggcctgat	atggccacaa	2820
20	gggtccctag	aagtgtgaga	gggaggcagg	ggagagtcag	agaggggacg	tgagaaggac	2880
	cactggccac	tgctggcttt	gagatggagg	aggggggtcc	cagccaagga	atgggggcag	2940
	ccgtcccatg	ctggaaaagc	aagcaatcct	ccccggtcct	gagggcacac	ggccctgccc	3000
	acgcctcgat	ttcaggccag	tgggacctgt	ttcagctttc	cggcctccag	agctgtaaga	3060
	tgatgcgttt	gtgttcagcc	actaagctgc	agtgattcgt	cacagcagca	aatggaatag	3120
25	cagtacaggg	aaatgaatac	agggacagtt	ctcagagtga	ctctcagccc	acccttggg	3179

001260: 9723560